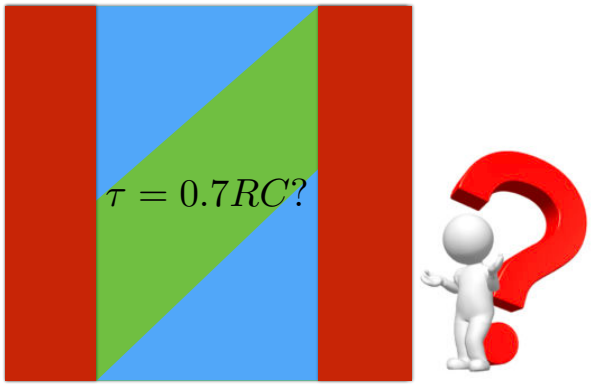
CHALMERS UNIVERSITY OF TECHNOLOGY



MCC092 Exercises 2018

Lena Peterson and Kjell Jeppson

Preliminary version
is being updated during course
Revised 0 times so far.

September 4, 2018

Contents

Chapter 1

Introduction

This is a set of exercises for the course MCC092. For the 2017 instance of the course we compiled problems previously used in exams and various sets of exercises through the years into one contained document. This document has further been updated for 2018. Still, it is not yet complete and will be updated some as we go. We have relied heavily on our old solutions from previous years. However, over the years our practice in how to express things in the course has evolved. So quite likely there are still quite a few errors, some due to this evolution and some just plain typos. We would be extremely happy if you students would report any errors you find to us. We promise to mention you in the list of contributors.

Chapter 2

Background material

This part is still missing, but it would be very good to have.

Here we have complied some problems on topics we expect you to have seen before. However, it may have been a long time ago so a brush-up may be good.

2.1 Logic

2.2 Capacitance and charge

2.3 Power and energy

Chapter 3

Logic functions, static CMOS gates, ILAs

3.1 Realizing logical functions

Exercise 3.1: Task tests understanding the implementation of static CMOS gates from logical expression. *Solution on page 43.*

Design **static CMOS gates** realizing the following Boolean expressions:

From old book. Check the homework exercises - are any of these included there? Missing solution.

a) $z = \overline{A \cdot B \cdot C \cdot D}$

b) $z = \overline{A \cdot B \cdot C + D}$

c) $z = \overline{(A + B + C) \cdot D}$

d) $z = \overline{A \cdot B + C \cdot (A + B)}$

Exercise 3.2: Task tests understanding the implementation of static CMOS gates from logical expression and the difference between implementing with given gates and as compound gates. *Solution on page 43.*

From old book. Missing solution.

Realize the following Boolean expressions using:

1. NAND, NOR, and-or-invert (AOI) and/or or-and-invert (OAI) gates.
2. Compound gates (that is implement one static CMOS gate that implements the full function.)

The functions are:

a) $z = A'B + AB'$ (XOR)

b) $z = AB + A'B'$ (XNOR)

c) $z = (AB'C' + A'BC' + A'B'C + A'B'C')$ which corresponds to the sum function, $SUM = A \oplus B \oplus C$, in a binary adder.

Which Solution more efficient in terms of the number of transistors used?

Exercise 3.3: BONUS Task tests understanding of pass transistor logic. *Solution on page 44.*

From old book. Missing solution.

Repeat exercise 3.2 but this time using MUXes, which can easily be implemented as pass-transistor logic.

Exercise 3.4: Task tests understanding of transistor scaling. Exam 2016-12-22 Task 1(a). *Solution on page 44.*

David Harris, the author of our textbook, holds quite a few patents. One of these is for static CMOS gates that have multiple outputs. The schematic for one such gate is shown in Figure 3.1, taken from the patent (but the gate in the figure is not one of the patented gates). What are the four logical functions for the four outputs Y1 through Y4 indicated in the schematic in Figure 3.1? (2 p)

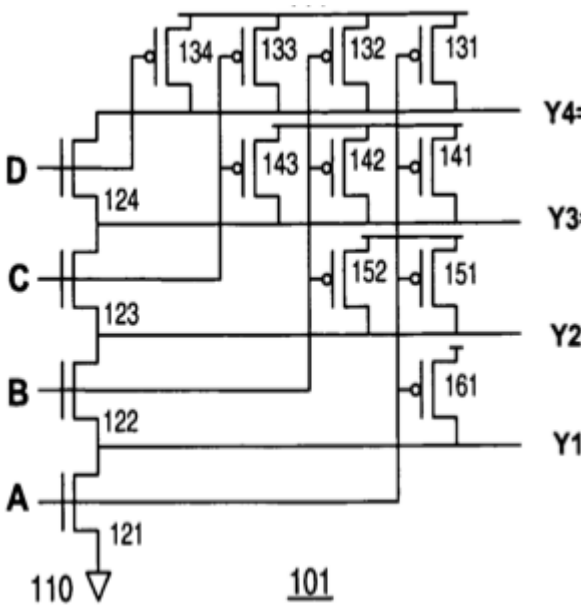


Figure 3.1: A schematic of the gate with multiple outputs taken from a patent application. V_{DD} is at the top of all pMOS transistors although is not so clear from the notation. There is no contact when lines cross.

3.2 Iterative logic arrays

Missing problems here. Add some that is not exactly like the lab.

Chapter 4

The MOS transistor

Exercise 4.1: Task tests understanding of MOS regions of operation. *Solution on page 44.*

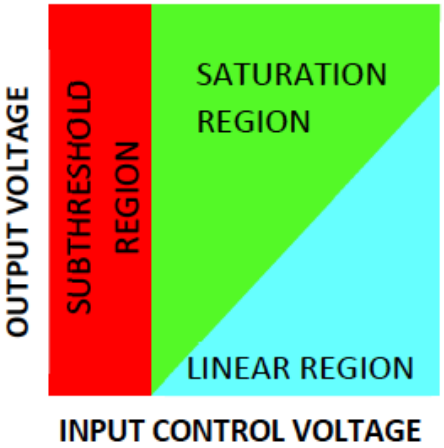


Figure 4.1: Regions of operation for a n-channel MOSFET.

a) Figure 4.1 shows the regions of operation for a n-channel MOSFET. What would be the equations for the borderlines between the different regions?

b) Draw a similar "regions of operation" diagram for a p-channel MOSFET.

Exercise 4.2: Task tests understanding of MOS model parameters. *Solution on page 45.*

The model parameters for an n-channel MOSFET, k_N and V_{TN}, in a certain MOSFET technology are given by k_N=900 µA/V^2 and threshold voltage V_{TN}=0.30 V.

a) Calculate the gate voltage overdrive V_{GT} if the supply voltage, V_{DD}, is 1.2 V and $V_{GS} = V_{DD}$.

b) Calculate the saturation current, I_{DSAT} when $V_{GS} = V_{DD}$.

c) Calculate the saturation voltage, V_{DSAT}.

Exercise 4.3: Exercise tests understanding of MOS capacitance . *Solution on page 45.*

a) Calculate the gate capacitance for a 1 mm wide MOSFET in the 65 nm CMOS process if its insulator capacitance per unit area, C_{ox}, is given as 20 fF/µm^2.

b) What would be the gate capacitance of the MOSFET in task 4.4b) if the effective gate length, L_{eff}, in the process is 45 nm and C_{ox} is 10 fF/µm^2?

c) Repeat task b) for the MOSFET width in task 4.4c).

a) Calculate the gate voltage overdrive V_{GT} if the supply voltage, V_{DD}, is 1.2 V and $V_{\text{GS}} = V_{\text{DD}}$.

b) Calculate the saturation current, I_{DSAT} when $V_{\text{GS}} = V_{\text{DD}}$.

c) Calculate the saturation voltage, V_{DSAT}.

Exercise 4.4: Exercise tests understanding of MOS effective resistance . *Solution on page 46.*

a) Calculate the effective resistances for two MOSFETs delivering maximum currents of 500 µA and 750 µA, respectively, at a supply voltage of 1 V.

b) If the effective resistance of a MOSFET in a certain technology is specified as 2 kΩ µm, what would be the effective resistance of a 5 µm wide MOSFET,

c) What if the MOSFET in b) was 280 nm wide? What would its effective resistance be then?

Chapter 5

The CMOS inverter

We use the CMOS inverter as a model for all static CMOS gates. Therefore it is essential to understand the inverter in detail.

5.1 Static characteristics

Exercise 5.1: Task tests understanding of voltage transfer curves, nMOS and pMOS transistors and and inverter bias point. *Solution on page 46.*

In Figure 5.1 you see the graphs of I_{DS} vs V_{OUT}, for the nMOS and pMOS FETs in a CMOS inverter. Match the three nMOS characteristics (a) through (c), shown to the left in in Figure 5.1. with the pMOS characteristics, (d) through (f) that corresponds the same input voltage. Then mark the corresponding bias points in the three voltage transfer curves (g) through (i) to the right. For more details on the red, blue, and green VTC diagrams refer to Figure 5.3.

Exercise 5.2: Task tests understanding of voltage transfer curves, nMOS and pMOS transistors and and inverter bias point. It is a variation of 5.1 *Solution on page 47.*

For three different input voltages, the output voltage of an inverter is swept from V_{SS} to V_{DD} while measuring the two MOSFET currents, I_{DSN} and I_{DSP}. The resulting current-voltage characteristics thus obtained are shown in Figure 5.2.

a) Match the two MOSFET currents for each of the three inverter input voltages, and find the bias points where the two currents are equal.

b) In the diagram of CMOS inverter regions shown in Figure 5.3 mark each of these three bias points with B, C, or D, depending on the region of operation in the V_{OUT} vs V_{IN} graph to which they belong.

Exercise 5.3: Task tests understanding of inverter switching voltage and VTC. *Solution on page 47.*

a) Add a secondary axis representing the "short-circuit" current through the inverter as shown in Figure 5.3. Sketch the "short-circuit" current through the inverter vs. V_{IN} based on your knowledge about the current through the current-limiting MOSFET.

b) Use the square-law MOSFET model and Kirchhoff's current law to derive the expression for the switching voltage of an electrically symmetrical CMOS inverter ($k_n = k_p$, $V_{TN} = -V_{TP}$)?

c) What happens to the switching voltage if $k_n = 4k_p$?

d) Derive an expression for the inverter switching voltage ($V_{IN} = V_{OUT}$) in the general case based on the square-law MOSFET models.

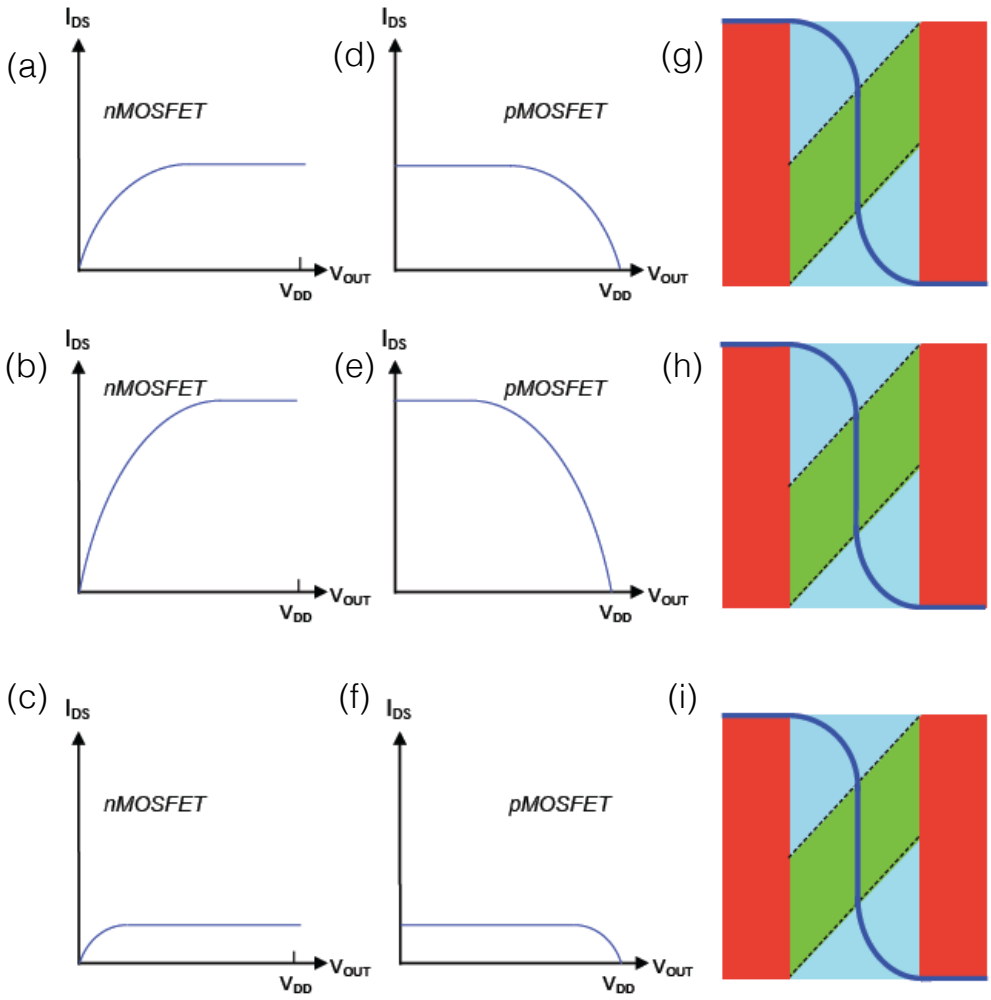


Figure 5.1: To the left, (a) to (c), are three nMOS characteristics each for one input voltage. In the middle, (d) to (f), are three pMOS characteristics for the same input voltages but not in the order that correspond to that of the nMOS characteristics. To the right, (g) through (i), are three transfer curves in which to indicate the bias points for cases (a) through (c).

Exercise 5.4: Task tests understanding of noise margins. *Solution on page 49.*

The solution for this problem is not included yet, because this problem is part of prelab 1.

To account for voltage fluctuations, i.e. noise, the valid high and low output voltages are usually defined within certain ranges like $0 \leq V_{OUT} \leq V_{OL,max}$, and $V_{OH,min} \leq V_{OUT} \leq V_{DD}$. Since CMOS is a robust technology, the input voltage can vary within ranges larger than those defined for valid output voltages without causing invalid output voltages, $0 \leq V_{IN} \leq V_{IL,max}$, and $V_{IH,min} \leq V_{IN} \leq V_{DD}$. These regions are usually defined from the two points, $(V_{OL,max}, V_{IH,min})$ and $(V_{OH,min}, V_{IL,max})$, on the VTC where the amplifications are equal to minus one, $A_v = -1$.

a) Derive expressions for the low and high noise margins, NML and NMH, as defined in the Figure 5.4 using the following expressions for $(V_{OL,max}, V_{IH,min})$ and $(V_{OH,min}, V_{IL,max})$:

$$\left(V_{OL,max} = \frac{V_{DD} + V_{TP} - V_{TN}}{8}\ ,\ V_{IH,min} = V_{SW} + \frac{V_{DD} + V_{TP} - V_{TN}}{8}\right)$$
$$\left(V_{OH,min} = V_{DD} - \frac{V_{DD} + V_{TP} - V_{TN}}{8}\ ,\ V_{IL,max} = V_{SW} - \frac{V_{DD} + V_{TP} - V_{TN}}{8}\right).$$

b) What are the explicit noise margin values in terms of fraction of V_{DD} if $V_{TN} = -V_{TP} = V_{DD}/5$?

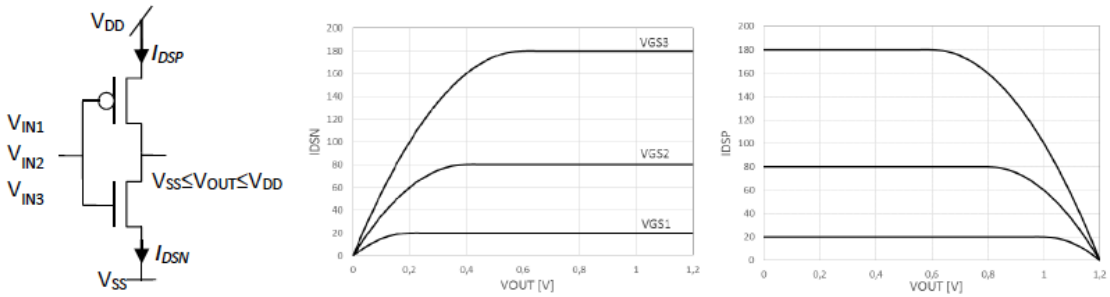


Figure 5.2: The CMOS inverter with three input voltages and corresponding curves for NMOS and PMOS transistors.

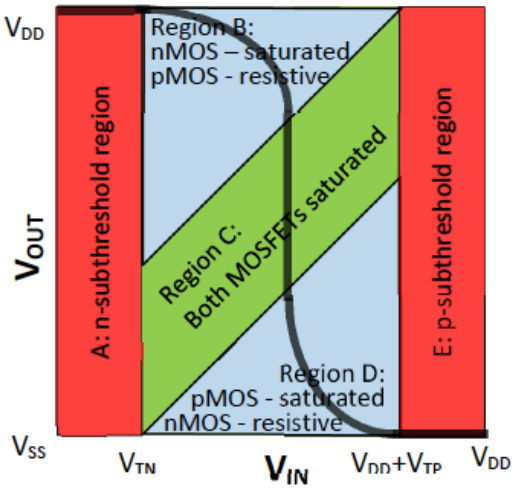


Figure 5.3: The CMOS inverter voltage transfer curve (VTC) with regions marked.

Exercise 5.5: Problem tests understanding of voltage transfer curves and noise margin. Slightly adapted from exam 2016-12-22 Tasks 2(a) and (c). *Solution on page 49.*

Figure 5.5 shows the voltage transfer curves for three of the four logical gates in the schematic Figure 3.1 simulated with all transistors having the same widths and with all four inputs, A-D, connected together, so that there is now only one input. The VTCs have been simulated in a DC analysis in Cadence in the usual 65 nm process. For simplicity you can assume that $V_{TN} \approx -V_{TP}$.

a) Which logical output of Y1-Y4 corresponds to each of the three VTCs named X (red), Y(green) and W(purple) in the graph in Figure 5.5? Motivate! (3 p)

b) In Figure 5.5 is also a plot of the derivative of the VTC for the X output (the red dash-dotted line). From the data given in Figure 5.5 calculate the noise margin for the X output. The graph in Figure 5.5 is repeated in larger scale in Figure A.1 for your convenience. (3 p)

5.2 Dynamic characteristics

Exercise 5.6: Task tests RC circuits. *Solution on page 49.*

From chapters. Missing solution.

Analyze the RC circuit and show that the time needed for the exponential decay of the voltage across the capacitor

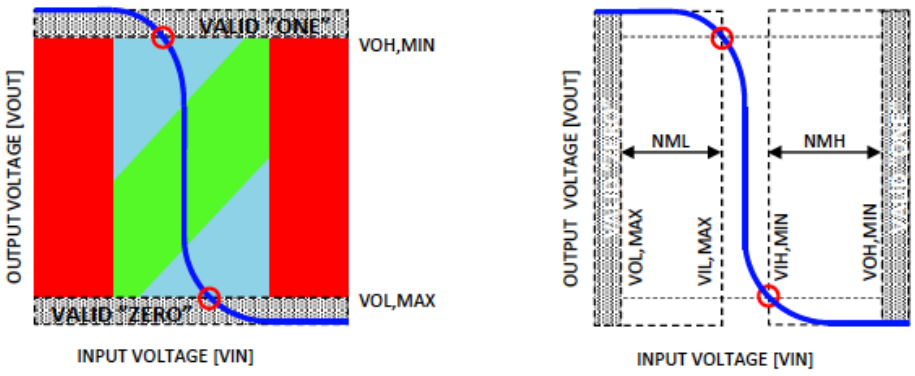


Figure 5.4: The The definition of noise margins.

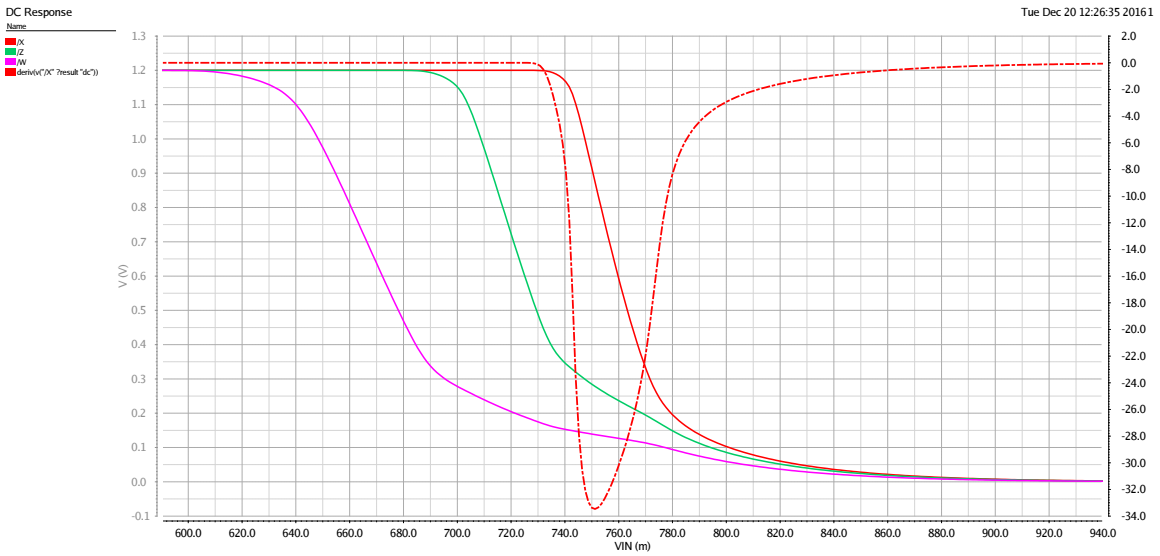


Figure 5.5: Voltage transfer curves (VTC) for the three outputs X, Y and W and the derivative of the VTC for output X.

to 50 % of the initial voltage, V_{DD}, is given by $t_d = RC\ln(2)$!

Exercise 5.7: Task tests understanding of inverter R,C and delay calculation. *Solution on page 49.*

a) What do we mean with an ideal inverter concerning its parasitic output capacitance?

b) Calculate the propagation delay of an ideal inverter driving an identical inverter! Assume the following MOSFET data: n-channel MOSFETs can sink 500 µA/µm channel width at $V_{DD} = 1$ V, and their input capacitances are 1.3 fF/µm. The p-channel MOSFET is made twice as wide as the n-channel device to obtain the same driving capability.

Exercise 5.8: Task tests understanding fanout-of-four delay. Exam 2012-10-26 Problem 2. *Solution on page 50.*

2017-09-07 Resistance values updated to reflect that 0.7 factor nowadays is not included in R.

a) Calculate the FO4 delay of a 0.35 µm CMOS process with $V_{DD} = 3.3$ V if the effective resistance in the timing model, R_{eff}, is 6 kΩ µm and the inverter input capacitance is 6 fF/µm. Assume $p_{inv} = 1$. (3 p)

b) What is the FO4 delay in a 65 nm process if we assume $V_{DD} = 1.2$ V, $R_{eff} = 2$ kΩ µm, $C_G = 4.5$ fF/µm, and $p_{inv} = 1/3$? (2 p)

Exercise 5.9: Task tests understanding of inverter R,C and delay calculation. *Solution on page 49.*

From chapters. Missing solution.

a) Assume that we, for simplicity, introduce a modified effective resistance $R' = R\ln(2)$, how large would this resistance be for the MOSFET in exercise 5.8?

b) How does the use of R' modify our delay model?

Exercise 5.10: Task tests understanding inverter operation, inverter switch voltage. Exam 2013-08-26 Problem 2. Slighly modified. *Solution on page 50.*

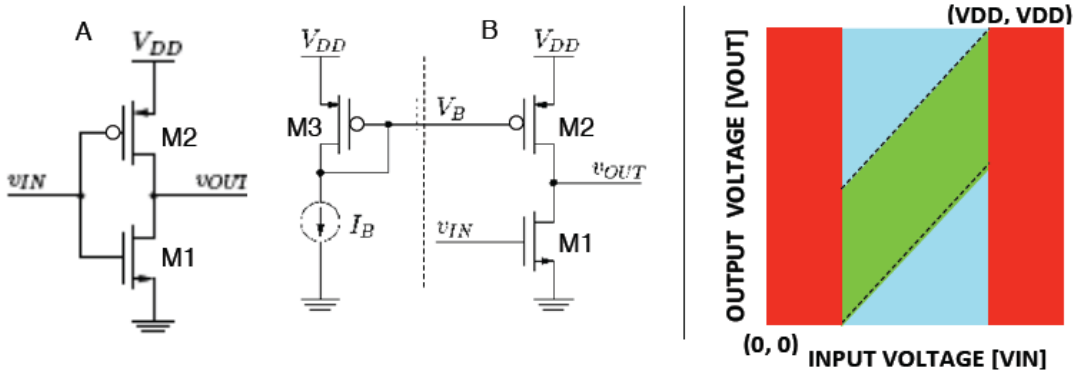


Figure 5.6: Two CMOS inverters. (A) is a regular CMOS inverter (B) is a pseudo-NMOS inverter intended for use as an amplifier. To the right is a diagram showing the MOSFET regions of operation in CMOS.

In Figure 5.6 you see the circuit diagram for two CMOS inverters, (A) and (B). Inverter (A) is an ordinary CMOS inverter that switches at an input voltage of $V_{DD}/2$. Inverter (B) is a pseudo-NMOS inverter intended for use as an amplifier. Its load p-channel MOSFET, M2, is biased at an unknown gate voltage V_B determined by a current mirror. A current mirror takes current I_B from a constant-current source and mirrors it to the inverter. Except for the biasing arrangement, the two CMOS inverters are identical. That is, transistors M1 and M2, respectively, are the same MOSFETs in both inverters. The rightmost diagram in the figure above shows the MOSFET regions of operation in CMOS.

a) Relate the two current gain factors k_1 and k_2 of MOSFETs M1 and M2 to each other considering that inverter (A) flips at $V_{DD}/2$ assuming symmetrical threshold voltages, $V_{TN} = -V_{TP} = V_{DD}/5$? (2 p)

b) Assuming $V_B = 0.6V_{DD}$, what is the switching voltage of the pseudo-NMOS inverter (B)? (2 p)

c) Calculate current I_B if $V_{DD} = 1.2$ V, and $k = 600$ mA/V^2! (2 p)

d) For what output voltage range are both MOSFET devices saturated in inverter (A)? Refer to the right-hand diagram showing the CMOS regions of MOSFET operation! (2 p)

e) For what output voltage range are both devices saturated in the pseudo-NMOS inverter? Refer to the right-hand diagram showing the CMOS regions of MOSFET operation! (2 p)

5.3 Tapered buffers etc.

Exercise 5.11: Task tests understanding of inverter delay calculation and minimization. *Solution on page 50.*

Mssing solution.

Four is sort of a magic number, if the number of loading inverters becomes much larger than four, it is often more efficient to insert an extra inverter with a better driving capability as a buffer between the original inverter and the capacitive load.

a) What driving capability should the inserted buffer inverter have to minimize the delay?

b) For what number of loading inverters does the inserted buffer shorten the propagation delay?

c) How does the parasitic output capacitance influence these critical numbers?

Exercise 5.12: Task tests understanding of inverter delay calculation and minimization. *Solution on page 51.*

Mssing solution.

For how big a capacitive load would the insertion of a non-inverting, two-inverter buffer give the shortest propagation delay?

Exercise 5.13: Task tests understanding of inverter delay calculation and minimization. *Solution on page 51.*

Mssing solution.

1. Determine the number of buffer inverters needed to minimize the delay if the load capacitance is 1000 times larger than the inverter input capacitance?
2. What would be the optimum tapering factor?

Exercise 5.14: Task tests understanding of how to deign a tapered buffer. Exam 2015-01-09 Problem X. *Solution on page 51.*

In a chip you are responsible for designing a driver for an output pad. The capacitance of the pad is (around) $1024C$ where C is the input of the minimum inverter in the process. The setup is shown in the figure below. The polarity of the output signal is not important in this particular case. Assume that the parasitic output capacitance is half of the inverter input capacitance (that is $p_{\text{inv}} = 0.5$). Tau in the particular process is 4 ps. (Tasks c) and d) of this exam problem will appear when we get to power and energy.)

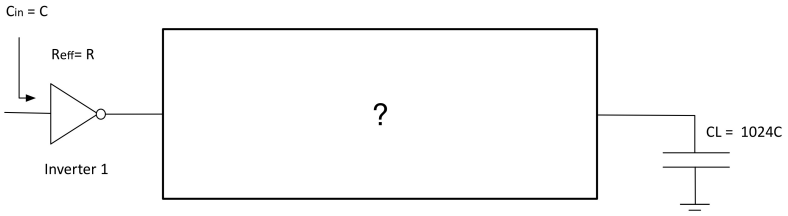


Figure 5.7: The setup for the driver that you are to design.

a) If minimum delay is the main design goal, how many inverters do you choose to use in the box with the question mark? How would you size them? Draw a figure of your entire inverter chain with the inverter sizes clearly marked. Motivate your design choices, but proofs are not required. (4 p)

b) For your design what is the delay? (2 p)

BONUS QUESTION

e) What if in a similar design situation as in teak a) the polarity at the output were of importance and the obvious design choice was an odd number of inverters? Would you add one more inverter or would you remove one? Discuss your considerations. (2 p)

Chapter 6

Delay for complex gates and paths

6.1 Gate delay

Exercise 6.1: Task tests understanding of calculating logical effort, g, and parasitic delay, p, for simple gates. *Solution on page 51.*

Imagine that you have started working with a CMOS process where for the pMOS transistors the maximum saturation current is only 1/3 of that for an nMOS transistor of the same width (as usual we assume that we use the minimum length for all transistors).

a) Find the parasitic delay, p, and the logical effort, g, for a 2-input NAND and a 2-input NOR gate in this process. Assume that p for the inverter (sometimes called p_{inv}) is 0.5.

b) Imagine that the NAND2 gate in this process, which you analysed in task a), is connected to one 3-to-1 scaled inverter at its output to create an AND gate. If the pMOS transistors in the inverter gate are twice as wide as the ones in the NAND2 gate, what the normalized delay of the NAND2 gate?

Exercise 6.2: Task tests understanding of calculating logical effort, g, and parasitic delay p, for complex gates. *Solution on page 53.*

In Figure 6.1 you see two complex gates that we have found in the textbook. Again the task is to calculate the parasitic delay, p, and the logical effort, g. Note that in these gates the logical effort will not be the same for all the inputs, so you have to calculate one g value per input. In contrast, the parasitic delay, p, is related to the output, so there cannot be more than one value for$p$ for one specific circuit topology. As usual, assume that the p-transistor current is twice the n-transistor current for the same transistor width.

Exercise 6.3: Task tests understanding of transistor scaling for same worst-case resistance. Exam 2016-12-22 Task 1(c). *Solution on page 53.*

Refer to Figure 3.1. What if your task was to ensure that the worst-case resistance is the same for the p-net and the n-net for all four logic functions, Y1 – Y4. How would you size the transistors then? Assume that the drive strength of an nMOS transistor is twice that of a pMOS transistor with the same width. There are multiple solutions – you only have to give one. (3 p)

6.2 Path delay

Exercise 6.4: Problem tests understanding of how to calculate the path delay. *Solution on page 53.*

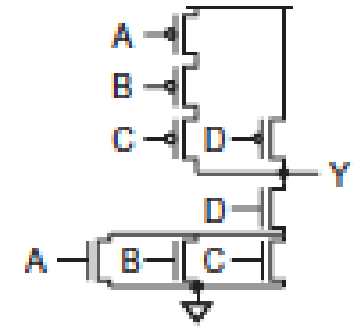
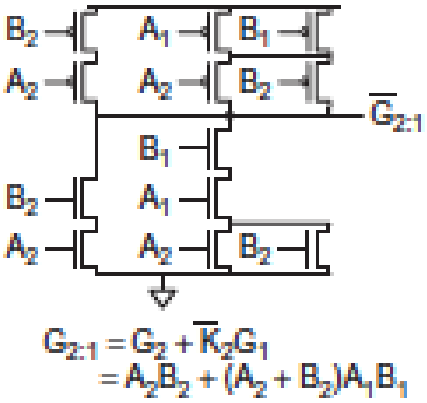


FIGURE 1.19
CMOS compound gate for function $Y = \overline{(A + B + C) \cdot D}$

(a) The first complex gate.



(b) The second complex gate.

Figure 6.1: Two complex gates for which to calculate logical effort and parasitic delay. Both are taken from the textbook.

In Figure 6.2 you see a path through a circuit made up of NAND and NOR gates. Add the missing data for the 3-input NAND gate and calculate the delay from A to B. In this process p_{inv} is 1.

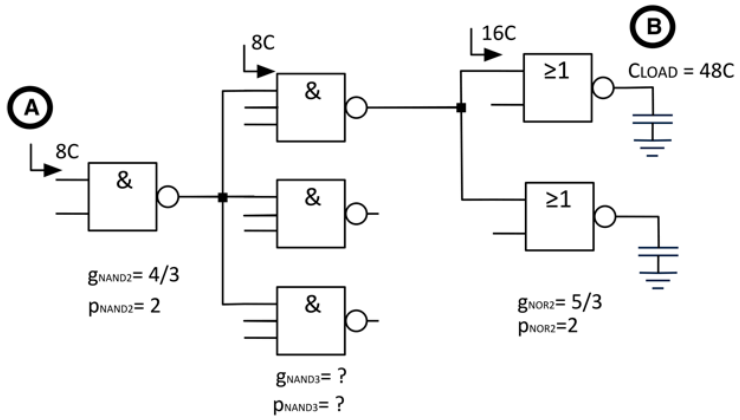


Figure 6.2: A path of NAND and NOR gates for which to calculate the path delay.

Exercise 6.5: Problem tests understanding of how to optimize the path delay. *Solution on page 54.*

Again consider the path shown in Figure 6.2. However this time assume that the load capacitance is 45C. Your task is to find the optimal sizes for the 3-input NAND and 2-input NOR gates to minimize the delay. What are the sizes and what is the delay when these sizes are applied?

Exercise 6.6: These three tasks test understanding of logical effort and path effort, and delay minimization Exam 2016-12-22 Problem 3. *Solution on page 55.*

In Figure 6.3 you see a block diagram schematic of a register file with 16 32-bit words and a 4-to-16 decoder that selects one of the 16 registers according to the address A[3:0].

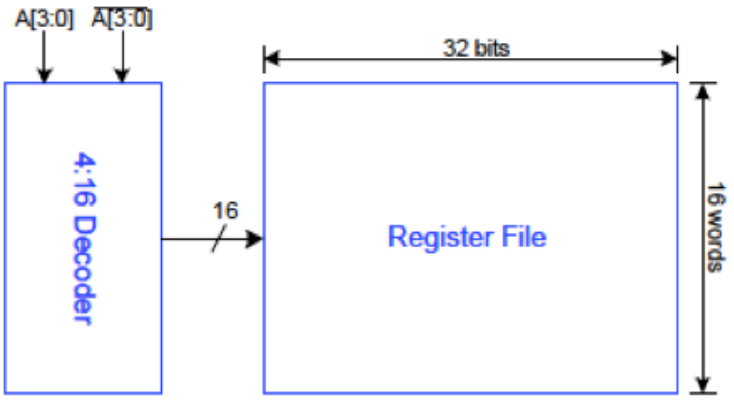


Figure 6.3: Block diagram of 4-to-16 decoder for addressing register file.

In this problem your task is to size the decoder circuitry shown in detail in Figure 6.4.

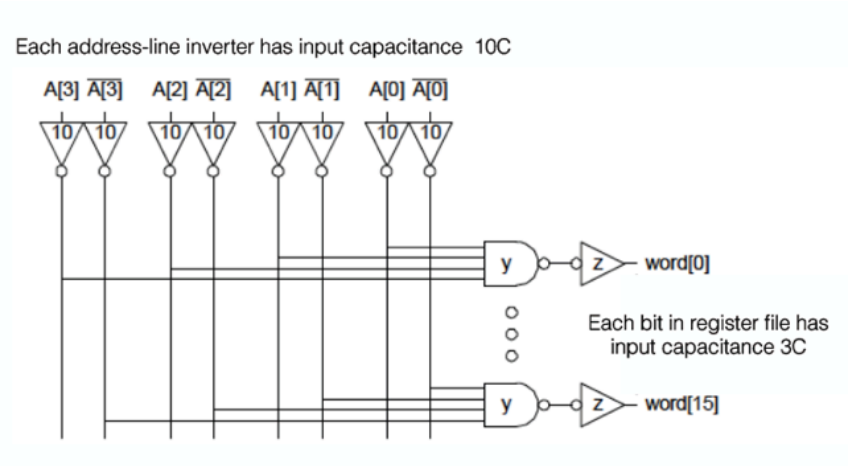


Figure 6.4: Detail of of 4-to-16 decoder for addressing register file.

a) How should the 4-input nand gate (labelled y in Figure 6.4) and inverter (labelled z in Figure 6.4) be sized for minimum delay with the assumptions given in Figure 6.4? Assume that an nMOS transistor has twice the current of a pMOS transistor of the same width. (5 p)

b) What is the resulting delay, including parasitic delays with the sizing from your result in task a)? Assume that the inverter's output capacitance is the same as its input capacitance. (2 p)

c) What if we had a wider register file of 16 64-bit words? Would it be faster to use two inverters in place of inverter z? (One would also have invert the address bits of course, but that could easily be achieved by swapping the lines for each address bit and its inverse). Motivate your reply. (3 p)

Chapter 7

Wire delay

Exercise 7.1: Task tests understanding of wire approximation. *Solution on page 58.*

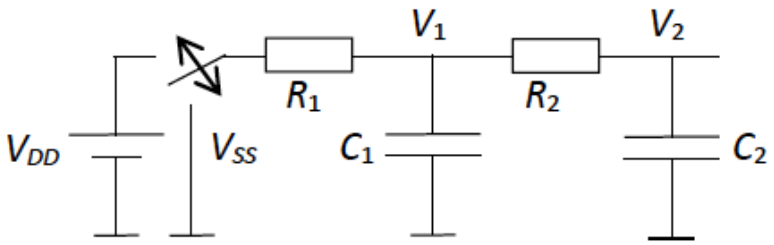


Figure 7.1: A two-node RC system.

Solution to be added for b and c

Shown in Figure 7.1 is a simplified version of the two-pole network from the book chapter.

a) Write down the two nodal equations for $V_1(t)$ and $V_2(t)$.

b) Find a method to convert these two nodal equations into second-order linear differential equation for $V_1(t)$ and $V_2(t)$.

c) Use the characteristic equation you found in task b), and identify $(s+a)(s+b) = s^2 + (a+b)s + ab = 0$ to find a simple way of determining the dominating time constant $1/a$ if $1/a \gg 1/b$.

Exercise 7.2: Task tests ability to calculate wire parameters. *Solution on page 56.*

In a certain CMOS process the wire sheet resistance is $0.2\,\Omega/\square$ and the wire capacitance is $0.4\,\text{fF}/\mu\text{m}^2$.

a) For a 200 nm wide wire calculate the resistance and capacitance for a wire that is 25 µm long.

b) Calculate the critical wire length for a wire that is 100 nm wide when the wire is driven by an inverter (that is, a repeater) with time constant $t_{rep} = 4.6$ ps.

Exercise 7.3: Task tests ability to calculate wire parameters. *Solution on page 56.*

In another CMOS process, the wire fringing field capacitance along the wire sidewalls cannot be neglected. This capacitance is $35\,\text{aF}/\mu\text{m}$ (including both sidewalls). The bottom-plate capacitance is $30\,\text{aF}/\mu\text{m}^2$. The wire sheet resistance is $0.10\,\Omega/\square$.

a) Calculate the wire resistance and wire capacitance for a wire that is 10 mm long and1 µm wide.

b) Calculate the delay from the input of a driver inverter to input of an identical receiver inverter a the other end of the 10 mm wire from task a), if the inverter can deliver 600 µA at V_{DD} = 1.2 V and its input and output capacitances are both 3.25 fF.

c) For this process and wire width, what is the critical wire length?

Exercise 7.4: Task tests understanding inverter and wire delay calculations. Adapted from exam problem from 2008. *Solution on page 56.*

An inverter is driving another identical inverter across a rather long RC wire. The inverter input and output capacitances are both C, and its resistance is R. The wire resistance is 4R and the wire capacitance 8C.

a) Draw a circuit diagram for the two inverters and the wire using a suitable model for the wire.

b) Use the diagram to calculate the wire RC delay from the output of the driver inverter to the input of the receiver inverter.

c) What if the *receiver* inverter were a NAND2 gate with input capacitance 2C. What would the change in the delay be then?

d) What if, instead, the *driver* inverter were a NAND2 gate with input capacitance 2C. How would the delay change then?

e) What if, instead, a branch wire (R, 2C) were added at the midpoint of the wire. How much would the delay increase then?

f) If the inverters in the original setup (task b) were properly scaled for minimum delay, what would be the resulting delay then?

Exercise 7.5: Task tests calculating wire and gate delay, logical effort. Exam 2016-08-22 Problem 3. *Solution on page 58.*

Figure 7.2 shows part of a clock-distribution network that comprises an inverter that acts as the clock driver, some wiring, and three identical NAND gates that act as clock gaters.

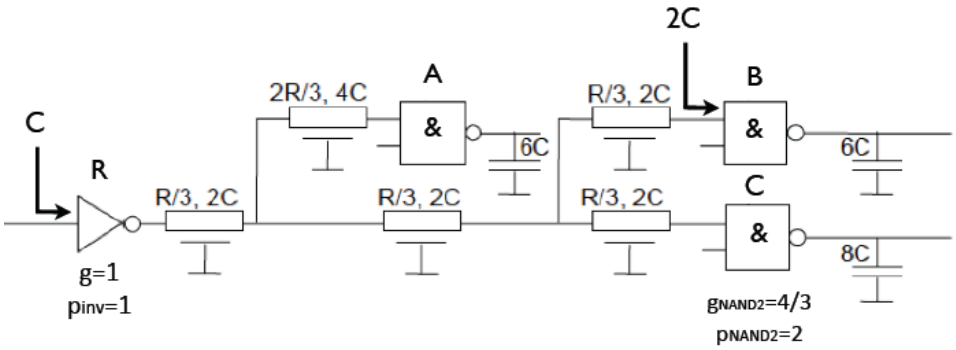


Figure 7.2: Clock-distribution network with a driver driving three receivers over different wires.

a) Calculate the clock skews at the inputs of all three clock gaters: A, B, and C. The clock driver has a driver resistance R and an input capacitance C. The identical NAND2 gates all have an input capacitance of 2C. (6 p)

b) Calculate the clock skews at the outputs of the three NAND2 gates: A, B and C. (4 p)

Exercise 7.6: Task tests calculating wire delay and sizing path to drive wire. Exam 2015-10-29 Problem 4. *Solution on page 59.*

Figure 7.3 shows part of a clock-gated network for distributing a system clock on a chip. The largest available inverter in the cell library (X200, C_{IN} = 72 fF, R = 100 Ω) was chosen to drive the wire network.

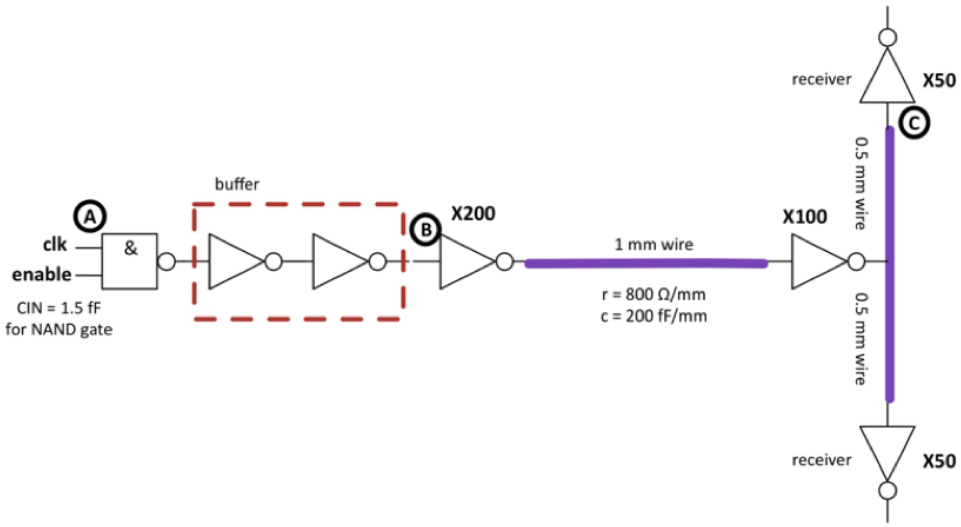


Figure 7.3: Clock-network for distribution of clock on chip.

a) Calculate the delay from the input of the X200 inverter (point B) to the input of one of the X50 receivers (point C). You may assume that the input of the X200 inverter is driven by an infinitely strong driver. (4 p)

b) There is a two-inverter buffer inserted between the clock-gating NAND gate and the X200 inverter. Design this buffer – that is, determine the sizes and/or input capacitances of the two buffer inverters for minimum delay from the CLK input of the NAND gate (point A) to the X200 inverter input (point B). (3 p)

c) Calculate the resulting delay from the CLK input of the NAND gate (point A) to the X200 inverter input (B). (3 p)

Exercise 7.7: Task tests calculating wire parameters and resulting delays. Exam 2014-08-25 Problem 2. *Solution on page 60.*

In Figure 7.4 you see the general layout of a static random-access memory (SRAM). The word lines (WL) select the particular word that is to be read or written. The bit lines (BL) carry the bit values out when reading and supply the bit values that are to be written when writing. The bit lines are routed in metal 1 (blue) and the word lines are routed in metal 2 (purple). In each memory cell the word line is connected to two minimum-size nMOS transistors for accessing that particular memory cell.

A word line is 0.1 µm wide, which is the minimum width for an M2 wire in this particular process. A word line has a capacitance of 0.1 fF/µm to ground and an inter-wire capacitance to one adjacent word line of 0.02 fF/µm and the M2 layer has a resistance of 0.1 Ω/□. A minimum-size nMOS transistor has a gate capacitance C_g = 0.1 fF. V_{DD} is 1 V.

a) Calculate the resistance for one WL.

b) Calculate the total capacitance for one WL including the capacitance of the access transistors.

c) Draw a circuit diagram for the WL and the inverter that is driving it. Assume the driver is an inverter with the same equivalent resistance as the WL resistance. Calculate the delay. For simplicity neglect the parasitic capacitance of the driver.

d) Calculate the energy required for accessing the memory when reading the memory once.

e) Estimate what would happen to the delay and energy computed in c) and d) if one could make the memory cells half as high and half as wide. Assume that the inter-wire capacitances are pure plate capacitances and the driver is re-sized so that its equivalent resistance still is the same as that of the wire. Reflect on the result.

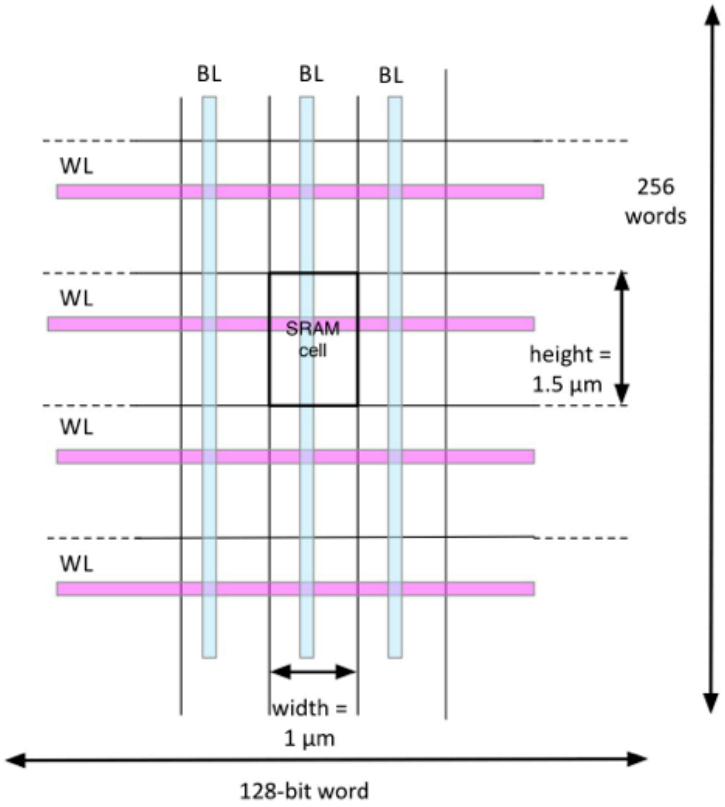


Figure 7.4: General layout of SRAM memory bank with 256 128-bit words. The word lines (WL) run horizontally in metal-2 (purple) and the bit lines (BL) vertically in metal-1 (blue).

Exercise 7.8: Task tests calculating wire delay and sizing inverter to drive wire. Related to lab 4. Exam 2016-12-22 Problem 5. *Solution on page 60.*

In Figure 7.5 you see a driver inverter loaded by four identical receiver inverters across an H-tree wire interconnect.

a) Calculate the FO4 delay for the driver inverter when loaded as shown in Figure 7.5. (5 p)

b) Determine the inverter resistance, R_{eff}, that minimizes the FO4 delay as calculated in a). You may assume that the inverter output capacitance, C_D, is equal to the inverter input capacitance, C_G. (5 p)

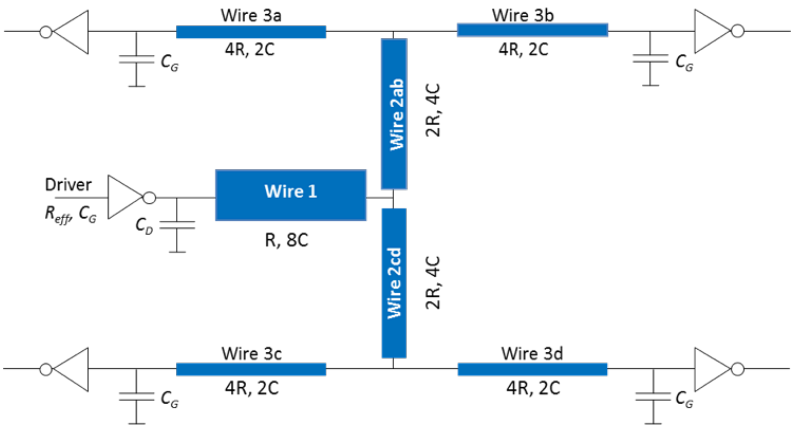


Figure 7.5: A driver inverter loaded by four identical receiver inverters across an H-tree wire interconnect.

Chapter 8

Layout

Exercise 8.1: Problem tests the understanding of continuous-line-of-diffusion layout. *Solution on page 61.*

a) Using Euler paths determine if it is possible to lay out the gate shown in Figure 6.1 a) with single-line-of-diffusion approach. If it is possible determine one order of the input signals in the layout that will work.

b) Using Euler paths determine if it is possible to lay out the gate shown in Figure 6.1 b) single-line-of-diffusion approach. Assume that the repeated signals, A2 and B2, will also be repeated in the diffusion line. If it is possible, determine one order of the input signals that will work.

Exercise 8.2: Problem test understanding of layout, parasitic delay and logical effort of multi-stage gates. Adapted from problem 1 in exam 2016-08-22. *Solution on page 62.*

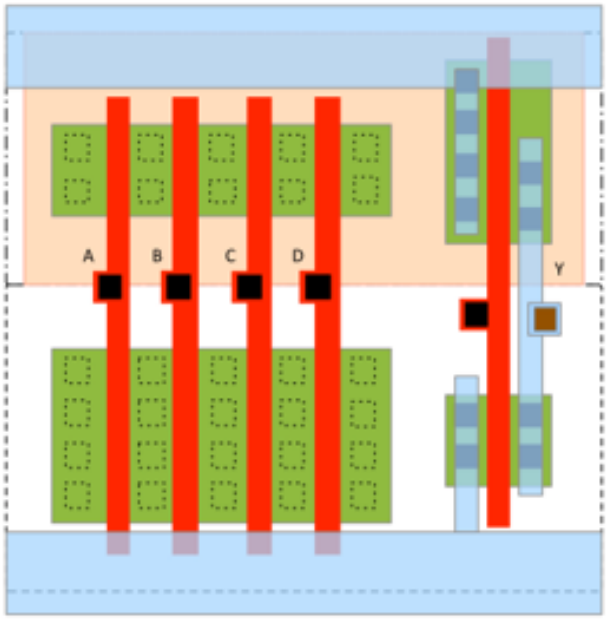


Figure 8.1: Template for 4-input AND gate.

For a 4-input NAND gate we have previously found that the logical efforts for all inputs are 2 and the parasitic delay is $4 \times p_{inv}$. In this problem you will layout and model a 4-input AND gate using such a NAND gate. You can assume that $p_{\text{inv}} = 1$.

a) In the cell layout template in Figure 8.1 you see the layout of one inverter to the right. To the left of that inverter is a continuous-line-of-diffusion template. Draw the layout for the 4-input NAND gate there; also

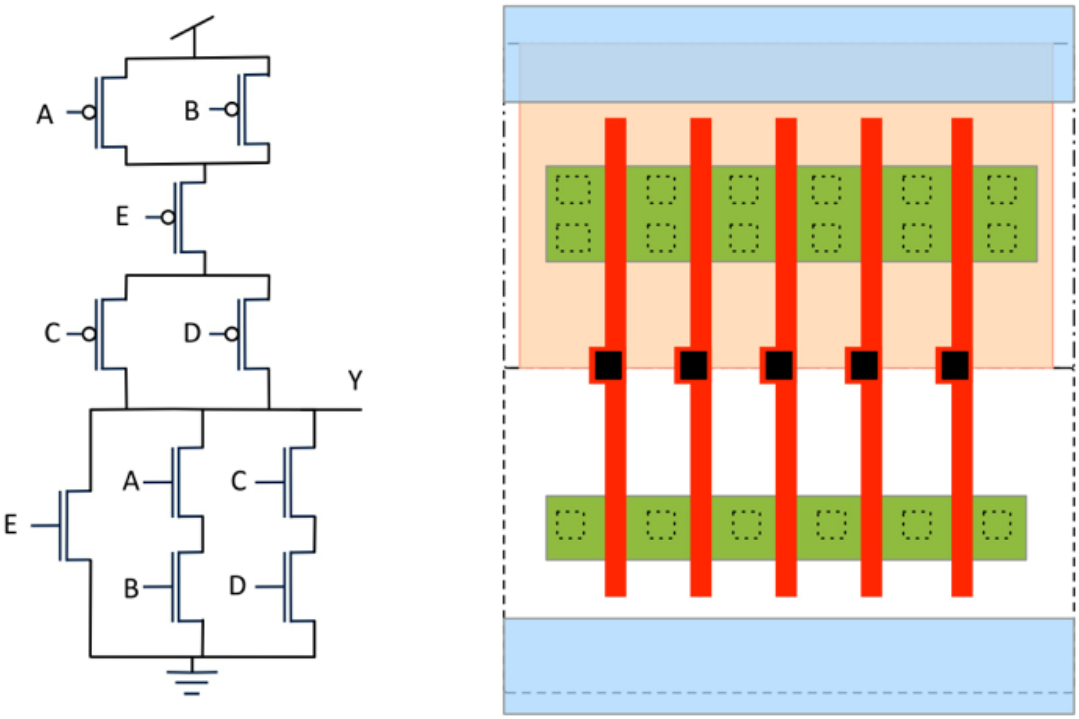
connect the output of the NAND gate to the inverter input, thus forming a 4-input AND gate. Draw the layout such that you minimize the number of diffusion areas connected to the output node of the NAND gate. (3p)

b) The parasitic delay, p, of a static CMOS gate is due to the capacitances of the diffusion areas connected to the gate output. Find the value for p_{NAND4} for your layout from task a). Assume that the capacitance of a diffusion area of a particular width is the same if it is shared between two transistors as if it is not shared. (2p)

c) For the 4-input AND gate, formed by the inverter and the 4-input NAND gate, find the logical effort, g_{AND4}, (the same for all four inputs), and parasitic delay p_{AND4}, for the entire gate. Use the p_{NAND4} value from task b) and find the relative transistor widths from the layout. (3p)

Exercise 8.3: Problem tests going from schematic to layout and from layout to schematic and logical function. Exam 2015-08-24 Task 1. *Solution on page 63.*

a) Figure 8.2a shows the circuit schematic for a compound gate that implements a 5-input logical function. Draw the corresponding layout in the template supplied in 8.2b . The layout should correspond exactly to the schematic – logical equivalence is not enough. Label all inputs and outputs. Indicate clearly any contacts to metal. (5 p)



(a) Schematic of the 5-input gate.

(b) A layout template in which to lay out the gate.

Figure 8.2: A 5-input static CMOS gate.

b) In Figure 8.3 is the layout for a compound static CMOS implementation of a 4-input logical function. Draw the circuit schematic for the gate and find the Boolean expression for the logical function. (5 p)

Exercise 8.4: Problem tests going from layout to schematic and identifying the logical function. Exam 2015-01-05 Problem 1. *Solution on page 63.*

Figure 8.4 shows the layout of another four-input standard cell.

a) Draw the corresponding transistor diagram. Make sure that your transistor diagram matches the layout exactly! (6 p)

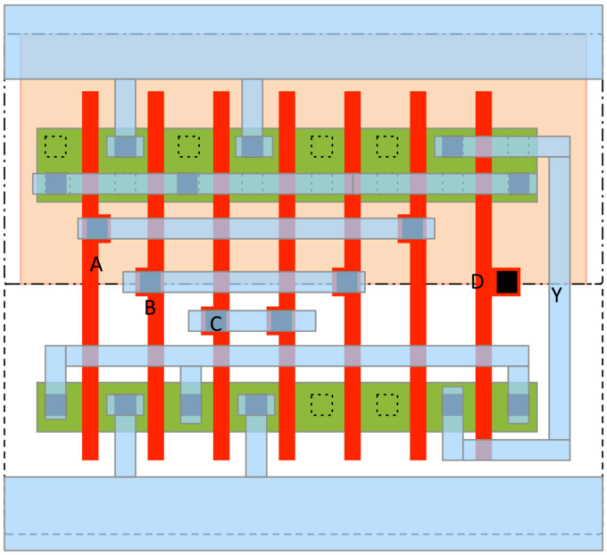


Figure 8.3: The layout for a 4-input static CMOS gate.

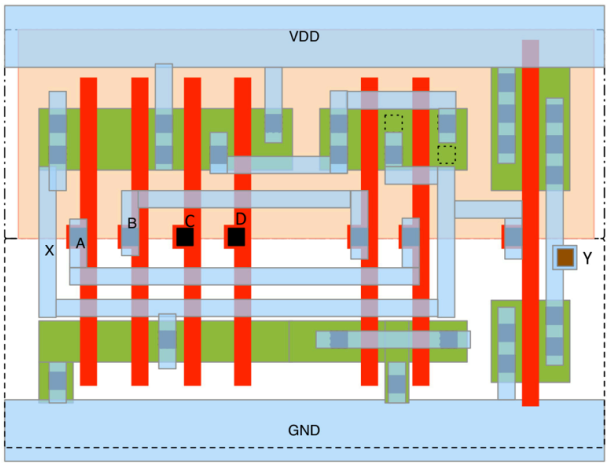


Figure 8.4: The layout for another 4-input standard cell.

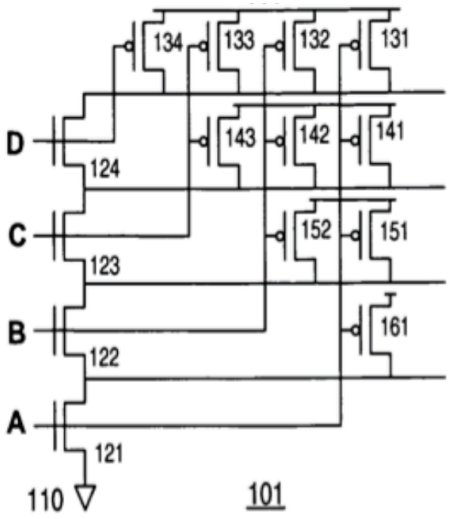
b) Identify the logical function that the layout implements. (4 p)

Exercise 8.5: Task tests going from schematic to layout. Exam 2016-12-22 Task 1(b). *Solution on page 64.*

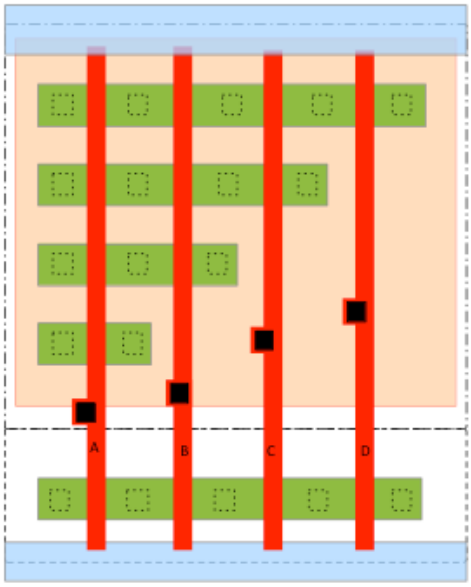
Draw the layout for the gate schematics in Figure 8.5a in the template provided in Fig. 8.5b. For simplicity we have assumed that all transistors have the same width although that may not be a good sizing. Hint: Remember that diffusion can be used to route V_{DD} or ground short distances. (5 p)

Exercise 8.6: Task tests understanding of what LVS does and reading layout. Exam 2013-10-22 Task 1. *Solution on page 64.*

What if we are confused by the error messages from the LVS, how shall we go about to find the discrepancies between the layout and the schematic entry? In other words, find the errors in the layout shown in Figure 8.6a for an AO22 gate. There are four discrepancies to detect! (10 p)

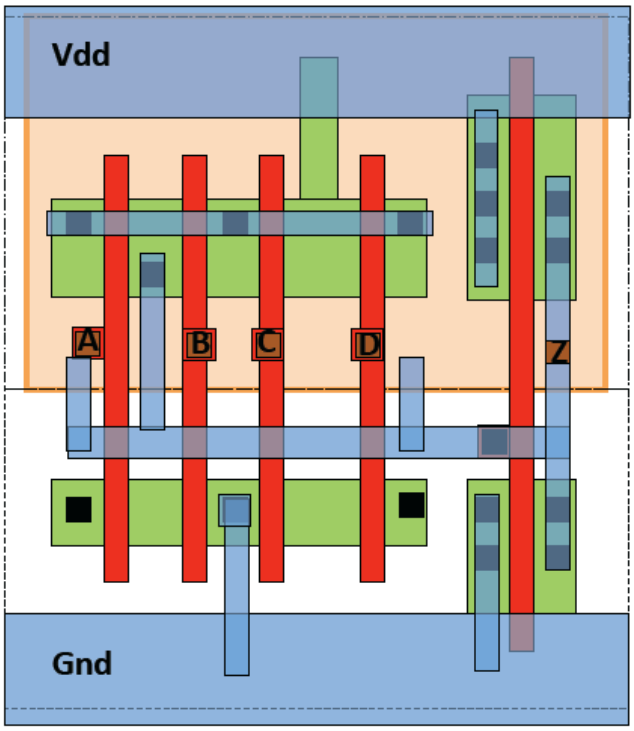


(a) A schematic of the gate taken from a patent application. V_{DD} is at the top of the schematic although is not clear from the notation.

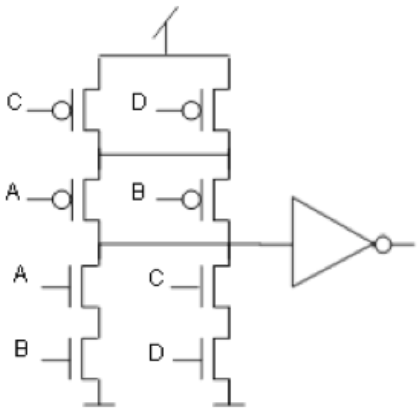


(b) A layout template in which to lay out the gate.

Figure 8.5: A static CMOS gate with four logical functions outputs.



(a) Layout of the AO22 gate



(b) Schematic for the AO22 gate.

Figure 8.6: An LVS problem - where in the layout are the discrepancies between the layout and the schematic?

Chapter 9

Sequential circuits

Exercise 9.1: Problem tests understanding of delay for ripple-carry adder together with flip-flops *Solution on page 65.*

You are to use an adder made up of your ripple-carry cell from prelab 2 and a sum cell in a pipelined processor. For simplicity we do not analyze the sum cell in detail; rather we assume the sum cell has a propagation delay of 60 ps. Also assume that the sum cell is of the type that takes the generated C_{out} from the same bit as one of its inputs, in addition to the input data bits A and B and C_{in}. (That is the sum cell shown in Figure 11.4 in Weste & Harris.)

Reminder in case you do not have it handy: The delay of the 8-bit ripple circuit in prelab 2 was around 340 ps with $p_{\text{inv}} = 0.8$.

You have available flip-flops with the characteristics given in Table 9.1

Table 9.1: Flip-flop timing characteristics

Flip-flop timing parameter	Value [ps]
Setup time, t_{setup}	50
clk-to-Q propagation delay, t_{pcq}	50
clk-to-Q contamination delay, t_{ccq}	35
Hold time,t_{hold}	10

a) Assuming the adder is the combinational circuit limiting the speed of you entire processor, what is the highest clock frequency with which you can clock it, if your adder has 16 bits?

b) Again assuming the adder is the combinational circuit limiting the delay of the processor, what is the highest clock frequency with which you can clock it, if your adder has 64 bits?

c) What if there is clock skew in you system? Assuming the maximum clock skew is 75 ps between any two flip-flops, how will the results in tasks a) and b) change?

Exercise 9.2: Problem tests understanding of delay for ripple-carry adder together with flip-flop hold violations. *Solution on page 65.*

Return to the setup in the previous exercise. Go back to the ripple-carry adder design based on the carry-chain you implemented in lab 2. Also assume that the sum cell has a contamination delay of 30 ps.

a) Estimate the contamination delay of the 16-bit ripple-carry adder designed from your ripple-carry chain. If necessary, you may assume that the inverses are also available from the flip-flips preceding the adder.

b) With these contamination delays, the flip-flop data in Table 9.1,the clock frequency calculated in task 9.1. a) and no clock skew determine if you have to worry about hold violations.

c) What if you have a clock skew of maximum 75 ps?

Exercise 9.3: Problem tests understanding of delay for adders together with flip-flops and critical path. This problem is from the exam 2016-08-22. It has been modified slightly to fit the purpose in this chapter. *Solution on page 65.*

In this course we have designed many adders but no multipliers. In this problem you will investigate how to use adders to implement binary multiplication and the performance of such an approach.

In Figure 9.1 is an example of a 6-bit binary multiplication from the Weste and Harris textbook.

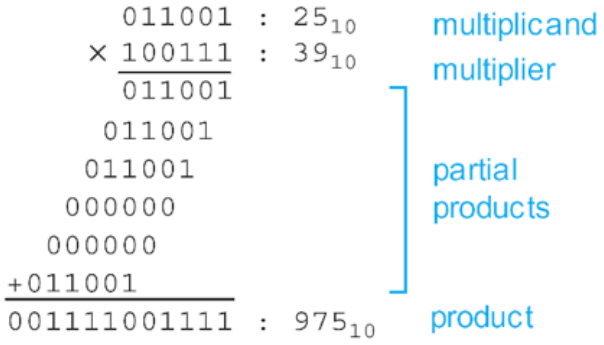
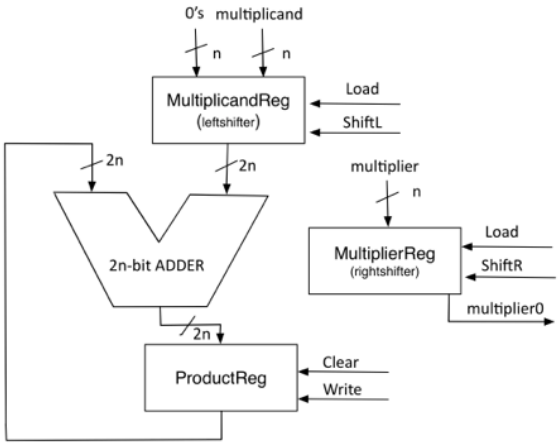
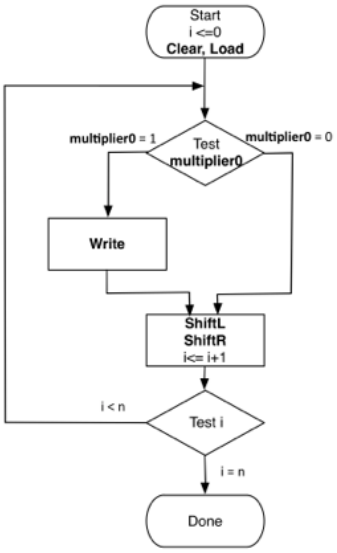


Figure 9.1: Example of a multiplication of two 6-bit binary numbers. From Weste and Harris textbook.

From the example it is clear that the partial products are just left-shifted versions of the multiplicand. Binary multiplication can thus be performed by repeatedly shifting the multiplicand to the left and adding it to the product. Figure 9.2 shows how a 2n-bit adder can be used to perform binary multiplication of two n-bit binary numbers. To the left in the figure you see the datapath with an adder, two shifters and a register, and to right the iterative control required to perform a multiplication.



(a) The datapath part of the multiplier.



(b) The control part of the multiplier.

Figure 9.2: An iterative multiplier that uses a 2n-bit adder.

In this problem, your task is to investigate the performance of this iterative multiplication for different types of adders and number of bits, n.

In Table 9.2 are worst-case propagation delays for two types of adders, for 8- and 16-bit additions.

Table 9.2: Adder worst-case propagation delays

Number of bits in adder n	Ripple-carry adder t_{pd} (ps)	Prefix adder t_{pd} (ps)
8	130	200
16	250	250

As you know for ripple-carry adders the worst-case delay grows linearly with the number of bits, n. We have not yet dealt with prefix adders, but their delay grows linearly with $\log_2(n)$.

Assume that the `ProductReg` is made up of flipflops with these characteristics: t_{setup} = 20 ps, t_{pcq} = 30 ps. For the shifter registers, assume there is a 30 ps delay from when the `ShiftR` and `ShiftL` signals are issued until the shifted output is available at their outputs.

For the control logic assume that each step takes one clock cycle. You may assume that the control signals that are the outputs from the control logic are perfectly synchronized with the clock.

a) Use the worst-case delay adder data in the table above to estimate the maximum clock frequency that can be used for the iterative 8-bit multiplier. With this clock frequency and assuming worst-case multiplier input data, how long would it take to complete one 8-bit multiplication? (4 p)

b) What if we extend the iterative multiplier from task a) to multiply two 32-bit binary numbers? How will its worst-case delay change? Assume that you can generate wider versions of the two types of adders in the table above. Which type of adder would you select? Motivate! For the selected type of adder, estimate the maximum clock frequency with which one could clock the multiplier control logic and still ensure a correct result. How long would it then take to complete one 32-bit multiplication with the worst-case multiplier input data? (6 p)

BONUS QUESTION

c) The proposed multiplier is not that well designed. Suggest one substantial improvement that could be made to the datapath and estimate how much that improvement would increase the maximum clock frequency calculated in task b). (4 p)

Exercise 9.4: Problem tests understanding of critical path, setup and hold violations. It is problem 5 from exam 2016-10-29. *Solution on page 66.*

Assume that you are designing an adder for the minimalistic 3-bit ArmStrong processor. The adder is built from three full adders such that the carry-out signal of the first adder is the carry-in signal to the second adder and the carry out from the second adder is the carry in of the third adder, as shown in Figure 9.3. At the input and output of the adder are two registers made up of flip-flops.

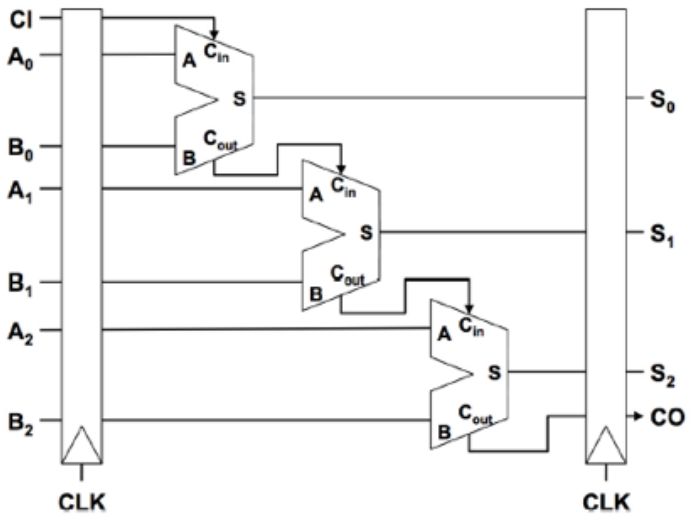


Figure 9.3: ArmStrong 3-bit adder.

a) If there is no clock skew, what is the maximum operating frequency of the circuit? Assume the delays for typical CMOS-process parameters, given in the leftmost column of values in Table 9.3. (2 p)

b) How much clock skew can the circuit tolerate before it might experience a hold violation? Again assume the delays for typical CMOS process parameters from Table 9.3. (2 p)

c) Assume we had characterized the flip-flop and full-adder cells also for the fast-fast and slow-slow process corners and measured the delays shown in the two right-hand columns in Table 9.3. Describe how you would go about extending the results from tasks a) and b) with these additional data so that you can be sure that your adder works correctly also for these two extreme corners. Carry out your proposed calculations. Did you have to modify your results from a) and b)? If so, what are the updated results? (6 p)

Table 9.3: ArmStrong timing characteristics

Delay	Typical corner measured value [ps]	Fast-fast corner measured value [ps]	Slow-slow corner measured value [ps]
Full adders			
t_{pd}, A or B → S	30	25	35
t_{cd}, A or B → S	22	16	20
t_{pd}, A or B → Cout	25	20	30
t_{cd}, A or B → Cout	22	17	25
t_{pd}, Cin → S or Cout	20	17	25
t_{cd}, Cin → S or Cout	15	12	20
Flip-flops			
t_{pcq}	35	28	40
t_{ccq}	21	16	24
t_{setup}	30	25	35
t_{hold}	10	5	20

Chapter 10

Power, energy and scaling

This chapter contains problems on power and energy. There is also a section on technology scaling in combination with delay, power, energy and area calculations, since the reason for technology scaling is often to decrease the power consumption while maintaining speed.

10.1 Power and energy

Exercise 10.1: Problem tests understanding of power. Problem is 5.2 in the textbook. *Solution on page 67.*

You are considering lowering V_{DD} to try to save power for a static CMOS gate. You will also scale the threshold voltages, V_T, proportionally to maintain speed (= performance). Will dynamic power consumption go up or down? Will static power consumption go up or down?

Exercise 10.2: Problem tests basic knowledge on how to calculate static and dynamic power dissipation. Exam 2014-10-31 Problem 1. Slighty revised. *Solution on page 67.*

The power consumption of a CMOS inverter can be minimized through circuit optimizations. How should each parameter in Table 10.1 be changed to reduce the three different components of the inverter power consumption? For each parameter indicate I for increase, D for decrease or N for does not affect this type of power consumption.

Table 10.1: Table for power optimization.

Type of power to minimize	Capacitive load C_L	Supply voltage V_{DD}	Treshhold voltages V_T	Transistor widths W
Dynamic power consumption due to the charging and discharging of the capacitive load C_L				
Power consumption due to shortcircuit current during transition (assuming a fixed rise and fail time at the inverter input)				
Static power dissipation, that is the power due to the transistor leakage currents				

Exercise 10.3: Problem tests understanding of subthreshold leakage. *Solution on page 67.*

Shown in Figure 10.1 are the static power entries for a 2-input NAND gate in the .lib file from a standard-cell library. When we inspect the power entries, we notice that one of the input combinations for this gate has a much lower leakage power than do the other three. Why is this? Explain!

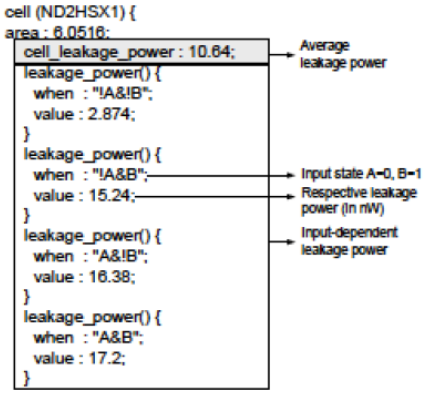


Figure 10.1: Data from lib file for 2-input NAND gate.

Will the ripple-carry gate that you have designed in labs 2 and 3 exhibit a similar leakage pattern, or not? That is, is there a best or worst combination of inputs when it comes to leakage? If so, what for input combination do we have this situation?

Exercise 10.4: Problem tests ability to calculate dynamic power dissipation and understanding of its origins. Exam 2015-01-05 Problem 3 c)-d). *Solution on page 68.*

Refer back to your solution to Exercise 5.14 or to its solution on page 51. Also, assume that the operating frequency is 200 MHz and that α is 0.25 for the signal that drives the pad. V_{DD} is 1 V.

c) For your design what is the dynamic power consumption? (2 p)

d) From the usual formula used for the dynamic power consumption is seems that to minimize the power consumption during switching there should be no inverters in the box in Figure 5.7. Why is this conclusion incorrect? (2 p)

Exercise 10.5: Problem tests ability to calculate static and dynamic power dissipation and knowledge of power and energy. It is adapted from two examples in the text book. Exam 15-10-29 Problem 3. *Solution on page 68.*

A digital system on a chip has 1 billion transistors. Of these 50 million are used in static CMOS logic gates and the rest are used in memories. The two parts of the chip are illustrated in Figure 10.3, which also includes relevant chip and process data.

a) Using data from Figure 10.3, estimate the power due to dynamic switching at a clock frequency of 1 GHz, if we neglect the effects of short-circuit current and wire capacitances. (2 p)

b) Using data from the figure above, estimate the static power consumption for the chip. (2 p)

c) What if the logic part can be redesigned so that fewer transistors than before require the low V_T: only 1 % of the original number, while the total number of logic transistors has to be increased by 20 %. The activity factor stays the same. What are the effects of this redesign on the static and dynamic power consumption for the logic part? (2 p)

d) If the measures above are not enough, one may have to use power gating to reduce the leakage further. What if we are to power-gate the entire logic part in the chip? We can tolerate only a 5 % drop of V_{DD} due to the resistance in the power switch, otherwise the increase in delay will be too large. How wide would our switch have to be if the pMOS transistor ON resistance is 2 k$\Omega\,\mu$m? Assume the dynamic power consumption of the logic calculated in task a). (2 p)

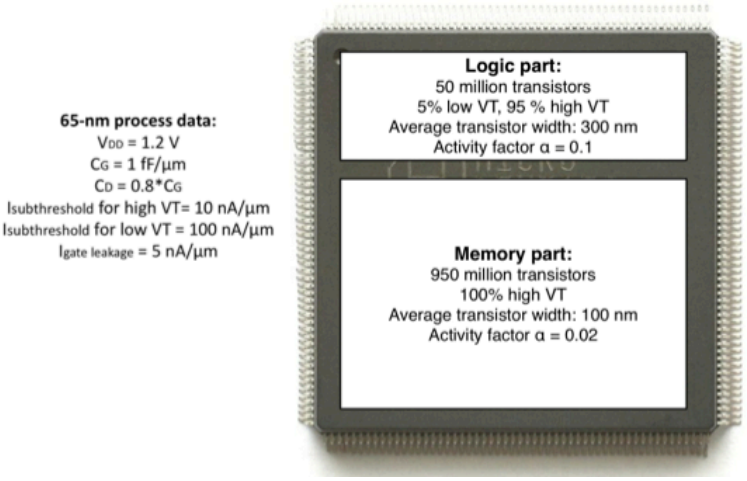


Figure 10.2: Data for the system on chip.

e) How much energy would be required to switch the power-gating transistor on and off once? How long a time with the leakage current for the logic part only, as calculated in task b), does that energy correspond to?(2 p)

BONUS QUESTION

f) When considering power gating, would it be better to start with the situation described in tasks a) and b) or the one described in task c)? Discuss the design considerations! (2 p)

Exercise 10.6: Problem tests ability to calculate static and dynamic power dissipation and knowledge of power and energy. It is adapted from two examples in the text book. Exam 15-10-29 Problem 3. *Solution on page 69.*

What if the processor in problem 10.5 is running at $V_{DD} = 1$ V instead. Would the width of the required switch, calculated in 10.5 d) change? What about the energy and time calculated in 10.5 e)?

Exercise 10.7: Problem tests static and dynamic power dissipation and their relation to V_{DD} and clock frequency. Exam 2016-12-22 Problem 4. *Solution on page 69.*

In this problem you will analyze the energy consumption for a video-rendering application run on a processor dedicated to this application. The processor can run at different supply voltages and has both a sleep mode and a hibernation mode. You final task is to determine if the hibernation mode is useful for this particular application.

Application: The digital video in the video-rendering application has 25 frames per second. One frame is 640x480 pixels. Each pixel is represented by 24 bits. The number of operations needed per pixel is 8 and these operations take 10 clock cycles to complete on the PP processor. The computations for one frame have to be completed in the time allotted for that frame.

PP processor: The PP processor used for this application has some characteristics shown in Table 10.2. You also know that it is fabricated in a CMOS process where the threshold voltages are 0.3 V. You may assume that the quadratic current equations hold in this process.

Sleep mode: The time it takes to enter sleep mode is 10 µs and it takes 20 µs for the processor to wake up from sleep mode. The energy required to switch the clocks off is 10 µJ.

Hibernation mode: The time it takes to enter hibernation mode is 1 ms and it takes 19 ms to wake up from hibernation mode. The energy required to turn off V_{DD} is 500 µJ.

a) Fill in the four empty cells in Table 10.2 above with reasonable values. (4 p)

b) Considering only dynamic power, how much energy will be used for one frame of the video-rendering application in these two cases: 1.2 V and 0.8 V supply voltage? (2 p)

Table 10.2: Data for the PP processor

Supply voltage V_{DD} [V]	Maximum clock frequency [GHz]	Current due to dynamic power consumption @ max clock frequency and a realistic activity factor [mA]	Idle current @ room temperature @ max clock frequency and a low activity factor [mA]	Static current in sleep mode @ room temperature (clock signal turned off for logic, but clock generation maintained) [mA]	Static current in hibernation mode (clock generation stopped and internal supply voltages turned off) [µA]
1.2	1.0	600	100	60	60
1.0	?	?	80	37.5	60
0.8	?	?	64	28	60

c) How much energy will be dissipated due to static power consumption for one frame for the two cases: 1.2 and 0.8 V supply voltage? Assume that the processor immediately enters sleep mode when no computations are required for the video-rendering application. (2 p)

d) What is your recommendation regarding the hibernation mode? Should it be used or not for the video-rendering application? Motivate your reply using data given in this problem and your results from tasks a)-c). (2 p)

Exercise 10.8: Problem tests ability to calculate delay with different threshold voltages. Exam 2013-10-21 Problem 5. *Solution on page 71.*

Need to add existing solution. Should this problem move, since it actually does not require any power calculation?

In a processor we have two circuit blocks, A and B, working in parallel. Both blocks are designed using the low-V_T standard cell library (where $V_{TL} = 0.2V_{DD}$). The delay of block B is only two thirds of the delay of block A, so in a way it is too fast and in some sense probably wasting energy. For block B the supply voltage could be reduced to save energy, or equivalently, the threshold voltage could be increased. The CMOS process used for the design comes with three different threshold voltages, low V_T, standard V_T and high V_T. By redesigning block B with cells from a library with a higher V_T, we can reduce the static leakage power while timing constraints are still met. Remember that a 100 mV increase in VT leads to a factor of ten reduction of the subthreshold current. The standard V_T: $V_{TS} = 0.3V_{DD}$ and the high V_T: $V_{TH} = 0.4V_{DD}$.

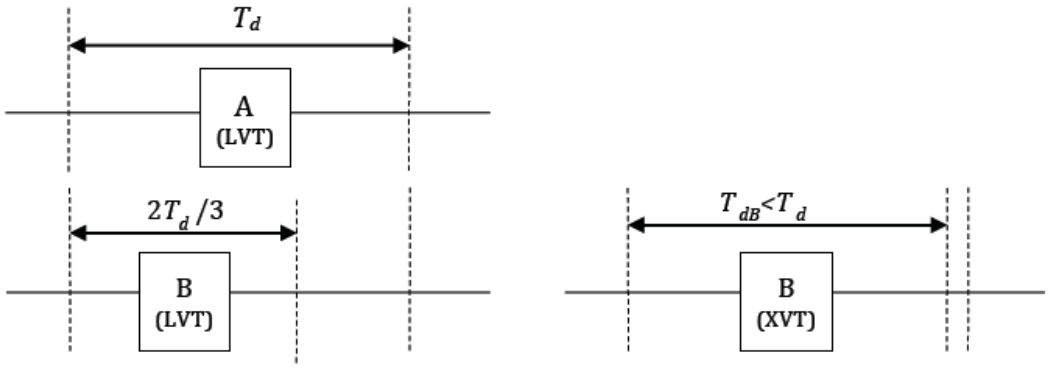


Figure 10.3: Illustration of tidming situation with blocks A and B. With the VLT library, the delay of block B is 2/3 of that of B. Which library can we select for B while still maintaining the situation that A determines the critical delay.

What if we were to redesign block B using a higher threshold voltage, which cell library would we use, the standard SVT library or the high HVT library, if we still were to meet the T_d timing constraint? Assume that the simple square-law MOSFET saturation current model can be used to estimate the MOSFET driving capability! All

parameters except the threshold voltages are the same among libraries.

10.2 Technology scaling and other scalings too

In this section we have collected problems related to technology scaling, but also problems that deal with other scalings, so that there is a mix.

Exercise 10.9: Problem tests ability to calculate scaling dependence for secondary parameters *Solution on page 71*.

Technology scaling is (or maybe, one now should say, was) performed to increase speed while lowering power consumption and increasing packing density, thus decreasing chip area, or enabling more functions on the same chip area. The classic scaling is the Dennard scaling. In table10.3 the scaling of the primary parameters is given at the top part of the table. Your task is to derive the scaling for each of the secondary parameters from the primary ones. It is best to do it in order since many of the derived parameters depend on each other.

Table 10.3: Table for Dennard scaling.

Parameter	Sensitivity expression	Dennard scaling, scaling factor S
Scaling parameters		
L: length		$1/S$
W: width		$1/S$
t_{ox}: gate oxide thickness		$1/S$
V_{DD}: power supply voltage		$1/S$
V_T : threshold voltage(s)		$1/S$
NA : substrate doping		S
Device characteristics		
β: current factor	$\frac{W}{L}\frac{1}{t_{ox}}$	
I_{DS}: transistor current	$\beta(V_{DD}-V_T)^2$	
R_{eff} resistance	$\frac{V_{DD}}{I_{DS}}$	
C: gate capacitance	$\frac{WL}{t_{ox}}$	
τ: gate delay	$R_{eff}C$	
f: clock frequency	$\frac{1}{\tau}$	
E: switching energy (per gate)	CV_{DD}^2	
P: switching power (per gate)	Ef	
A: area (per gate)	WL	
Switching power density	$\frac{P}{A}$	
Switching current density	$\frac{I_{DS}}{A}$	

Exercise 10.10: EXTRA Problem tests ability to calculate scaling dependence for secondary parameters *Solution on page 71.*

For quite a few years semiconductor manufacturers resisted scaling down the supply voltage and threshold voltages, since it would making interfacing with existing technologies harder. To see the effect of this type of scaling repeat exercise 10.9 but with "1" as the entries for V_{DD} and V_T.

Exercise 10.11: Problem tests ability to apply process scaling to delay and power. Exam 2011-08-24, task 4a. *Solution on page 71.*

The FO4 delay of the AMS 0.35 µm CMOS process running at 3.3 V is 125 ps. What would the FO4 delay be of a 0.13 µm CMOS process running at $V_{DD} = 1.8$ V?

Exercise 10.12: Problem tests ability to calculate static and dynamic power dissipation under the effect of technology scaling. Exam 2012-10-26 Problem 5. *Solution on page 72.*

The work horse product of the California-based microprocessor company MegaProcessor is a singlecore microprocessor manufactured in a 90 nm CMOS technology. The microprocessor operates at 3.8 GHz with a 1.2 V supply, a 100 W power dissipation, and a die area of 200 mm^2. The company is now about to design a dual-core microprocessor in the same technology, by duplicating the single-core design.

a) What would be the frequency and supply voltage for the dual-core design if the same size of the heat sink is to be maintained, that is if $P_{DUAL\_CORE} = P_{SINGLE\_CORE}$? Assume that 100 % of the power dissipation is due to dynamic power, and, for simplicity, assume that the frequency of operation is roughly linearly proportional to the supply voltage. (3 p)

b) When the single core design from above is moved to the65 nm technology node, what would be its size, power and frequency of operation with a 1 V power supply? (3 p)

c) Now consider the following what-if situation: assume that 10 % of the 100 W total power dissipated by the single-core in task a) is due to static (leakage) power and that 90 % is due to dynamic switching power. Furthermore, in the 65 nm process the standard threshold voltage (svt) is almost 50 mV lower than in the 90 nm process, yielding a four-fold increase in leakage power at room temperature. What would be the total power dissipation of the single-core processor when transferred to the 65 nm process? How many percent would be leakage power? (4 p)

Exercise 10.13: Problem tests ability to minimize dynamic power dissipation under constraint of constant delay. Exam 13-08-26 Problem 3. *Solution on page 72.*

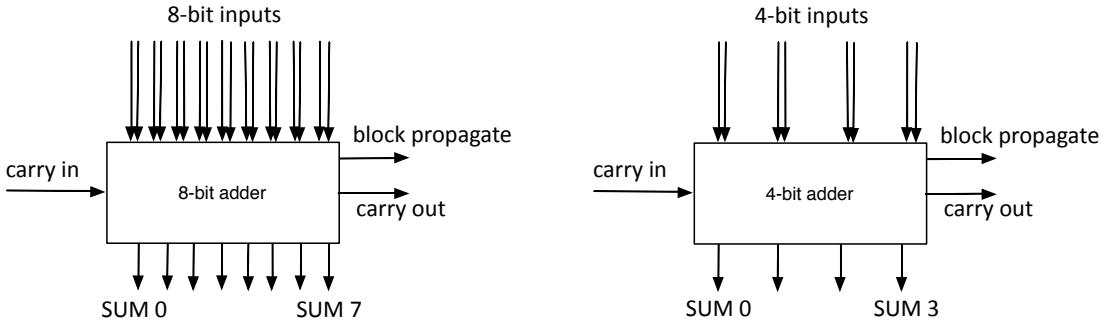
One reason for going to multi-core computer systems is to reduce the total power dissipation by lowering the supply voltage and then compensate for speed losses by using more than one core that work in parallel. A simplified assumption is that four cores at a fourth of full speed can do the same work as one core at full speed. From these simple assumptions and from the assumption that the delay is given by $0.7\frac{CV_{DD}}{I_{DSAT}}$, find out how much one can reduce the power dissipation by using a quad-core design, while still getting the same job done. Assume that we use the same CMOS process and the same core, but that an additional 20 % of the core capacitance must be added to the total chip capacitance for the control unit coordinating the four cores. Let us assume that V_T is 25 % of the original supply voltage. Use the simple square-law current model for the MOSFET saturation current!

Chapter 11

Adders

Need to add more adder problems.

Exercise 11.1: Problem test understanding of block-propagate signal and its use to create faster adders. Exam 2017-08-21 Problem 6. *Solution on page 73.*



(a) 8-bit adder.

(b) 4-bit adder.

Figure 11.1: Two adders. In (a) is an 8-bit adder and in (b) a 4-bit adder.

You have designed an 8-bit adder circuit with eight SUM outputs, one carry output and one block-propagate output, as shown in Figure 11.1 (a). The propagation delays expressed in unit delays are shown in Table 11.1. You have now now been asked to design a 32-bit adder with as short a propagation delay as possible. Available to you are AO, OA, AND and OR gates which each have a delay of 1 unit delay.

Table 11.1: Adder propagation delays

Propagation delay	8-bit adder [unit delays]	4-bit adder [unit delays]
From carry in to carry out	9	5
From any data bit to carry out	9	5
From any data bit to block propagate	4	2
From carry in to highest SUM output	8	4
From any data bit to highest SUM output	8	4

a) Draw a diagram of how you would construct a fast 32-bit adder from your 8-bit adder. In addition to the 32 SUM bits your 32-bit adder should also output the block-propagate and block-generate signals for the entire adder. (5 p)

b) Derive the propagation delay of your adder as drawn in task a). Assume that all inputs to your adder arrive simultaneously. (3 p)

c) What if you also had designed a similar 4-bit adder, as shown in Figure 11.1 (b), with propagation delays given in the last column of are shown in Table 11.1. Would it be better to use the 4-bit adder rather than the 8-bit adder in your 32-bit adder? Motivate by drawing a diagram and calculating the delay. (2 p)

Exercise 11.2: Problem tests understanding of prefix adder sum generation and critical path. Exam 2016-12-22 Problem 6. *Solution on page 73.*

Figure 11.2 shows the beginning of the design of an unknown prefix adder. As you can see from the figure, in this type of adder no dot-operator cell in the forward or backwards tree drives more than two other dot-operator cells. (The triangles in the diagram are buffers that you can ignore in this problem.)

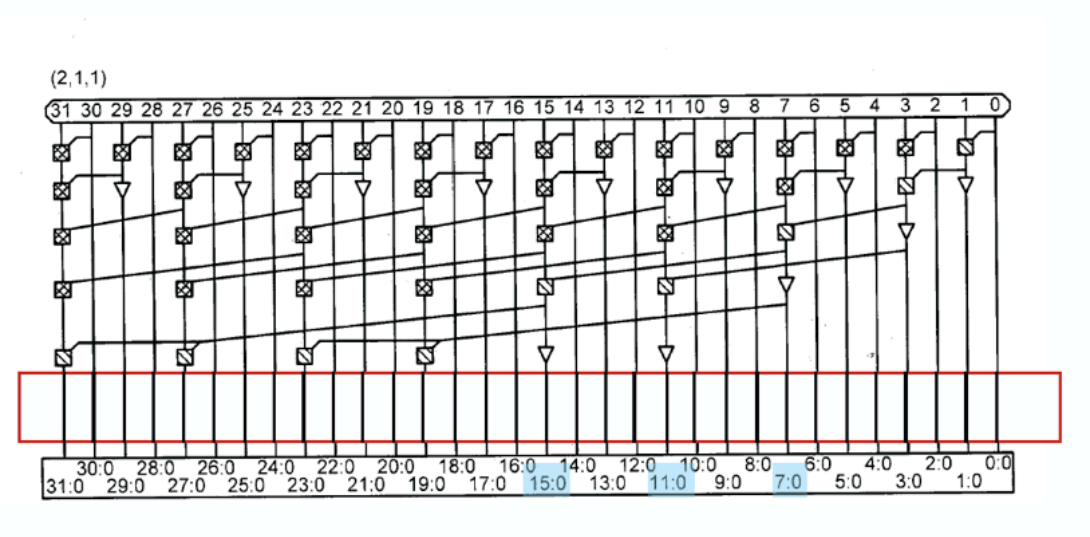


Figure 11.2: An unknown prefix adder where the output of any dot-operator cell is connected to a maximum of two other dot-operator cell inputs.

a) Your task is to complete the adder design by adding the missing dot-operator cells, in the red box, so that all input carries needed for the SUM operations are available at the bottom. As an example, group carries C15:0, C11:0 and C7:0 (indicated with blue in Figure 6) are already available to form SUM16, SUM12 and SUM8, respectively. A fully correct solution should maintain the design principle that no cell in the tree drives more than two other dot-operator cells. Solutions that do not fulfill this principle, but are logically correct, will give partial credit. The figure is repeated twice at the end of the exam so that you tear off and turn in with you solutions. (7 p)

b) What is the critical path of the prefix tree you have drawn in task a)? Indicate it clearly in the complete tree that you hand in. (3 p)

Chapter 12

Solutions

12.1 Introduction

12.2 Background material

12.3 Logic functions

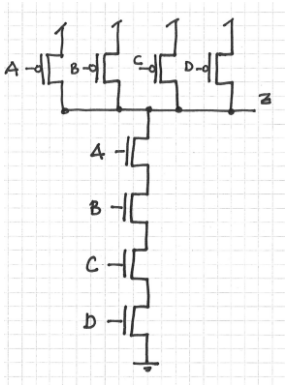
Solution 3.1 *Problem is on page 9.*

a) The circuit is a 4-input NAND gate. One solution with corresponding Karnaugh maps (ones and zeros) is shown in Figure 12.2. With the zeros covered it is easy to directly find the n-net, because the n-net implements the inverse of the logical function; in this case AND.

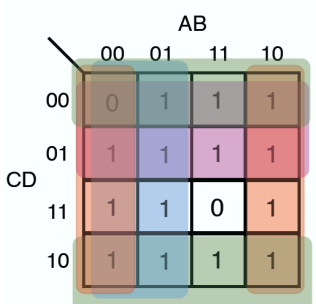
b) The logic gate is an AND-OR-INVERT (AOI) 3+1 gate. One solution with corresponding Karnaugh maps (ones and zeros) is shown in Figure **??**.

c) The logic gate is an OR-AND-INVERT (OAI) 3+1 gate. One solution with corresponding Karnaugh maps (ones and zeros) is shown in Figure **??**.

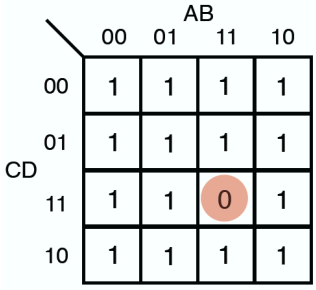
d) One solution with corresponding Karnaugh maps (ones and zeros) is shown in Figure **??**.



(a) Gate.



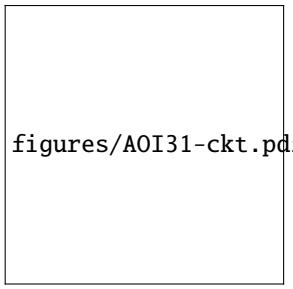
(b) Ones covered.



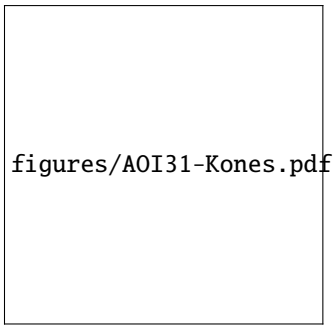
(c) Zero(s) covered.

Figure 12.1: NAND4 (a) gate and corresponding Karnaugh map with (b) ones and (c) zeros covered.

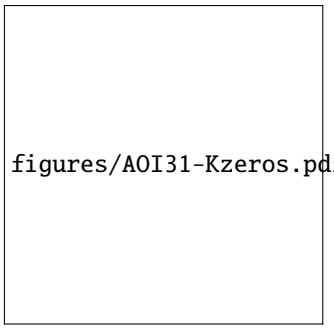
Solution 3.2 *Problem is on page 9.*



(a) Gate.



(b) Ones covered.



(c) Zero(s) covered.

Figure 12.2: NAND4 (a) gate and corresponding Karnaugh map with (b) ones and (c) zeros covered.

Solution 3.3 *Problem is on page 9.*

Solution 3.4 *Problem is on page 10.*

The logical functions are inverter (Y1) and nand (Y2-Y4). The expressions are found in Figure 12.3.

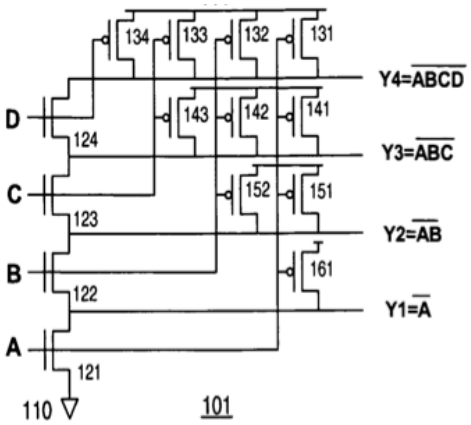


Figure 12.3: Schematic and logic functions for multi-output gate.

12.4 The MOS transistor

Solution 4.1 *Problem is on page 11.*

a) For the n-channel MOSFET we have the following conditions. For the transistor being ON:

$$V_{GS} \geq V_{TN} \tag{12.1}$$

For the transistor being in saturation:

$$V_{DS} \geq V_{GS} - V_{TN} \tag{12.2}$$

The borders between the regions is found when there is an equal sign in (12.1) and (12.2) instead of a greater-than-equal sign.

b) The corresponding conditions for the p-channel MOSFET are as follows below. For the transistor being ON it is

$$V_{GS} \leq V_{TP} \tag{12.3}$$

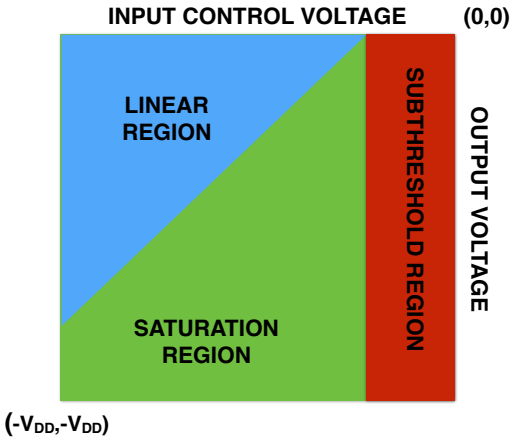


Figure 12.4: Regions of operation for a p-channel MOSFET. Note that all voltages are negative.

Remember here that V_{TP} is negative for the p-channel MOSFET. condition is

$$V_{DS} \leq V_{GS} - V_{TP}. \tag{12.4}$$

For the p-channel MOSFET the borders between the regions are found when there is an equal sign in (12.3) and (12.4) instead of a less-than-equal sign. So the equations for the borders are the same for the n-channel and the p-channel MOSFETs. The difference is that all voltages are negative for the pMOS case. The resulting diagram of the regions of operation for the p-channel MOSFET is shown in Figure 12.4.

Solution 4.2 *Problem is on page 11.*

a)

$$V_{GT} = V_{GS} - V_T = 1.2\,\text{V} - 0.3\,\text{V} = 0.9\,\text{V} \tag{12.5}$$

b) Assuming that the quadratic equation for the saturation current holds, we find:

$$I_{DSAT,\text{max}} = \frac{k_N}{2} V_{GT}^2 = \frac{900\,\mu\text{A/V}^2}{2} 0.9^2\,\text{V}^2 = 364.5\,\mu\text{A} \tag{12.6}$$

c) The saturation voltage is the drain-source voltage for which we have

$$V_{DSAT} = V_{GS} - V_T. \tag{12.7}$$

Thus in the situation in task b) we have

$$V_{DSAT} = 1.2\,\text{V} - 0.3\,\text{V} = 0.9\,\text{V}. \tag{12.8}$$

Solution 4.3 *Problem is on page 11.*

a) The expression for the gate capacitance is

$$C = C_{\text{ox}} \times L \times W \tag{12.9}$$

One would assume that in a 65 nm process the transistor length is 65 nm. However, it turns out that the physical length is somewhat less so we have to use 60 nm as the transistor length. So in this case we have:

$$C = 20\,\text{fF}/\mu\text{m}^2 \times 60\,\text{nm} \times 1\,\text{mm} = 20\,\text{fF}/\mu\text{m}^2 \times 0.060\,\mu\text{m} \times 1000\,\mu\text{m} = 1200\,\text{fF} = 1.2\,\text{pF} \tag{12.10}$$

b)

$$C = 10\,\text{fF}/\mu\text{m}^2 \times 45\,\text{nm} \times 5\,\mu\text{m} = 10\,\text{fF}/\mu\text{m}^2 \times 0.045\,\mu\text{m} \times 5\,\mu\text{m} = 2.25\,\text{fF} \tag{12.11}$$

c)

$$C = 10\,\text{fF}/\mu\text{m}^2 \times 45\,\text{nm} \times 280\,\text{nm} = 10\,\text{fF}/\mu\text{m}^2 \times 0.045\,\mu\text{m} \times 0.28\,\mu\text{m} = 0.13\,\text{fF} \tag{12.12}$$

Solution 4.4 *Problem is on page 12.*

a)

$$R_{\text{eff}} = \frac{1\,\text{V}}{500\,\mu\text{A}} = 2\,\text{k}\Omega \tag{12.13}$$

$$R_{\text{eff}} = \frac{1\,\text{V}}{750\,\mu\text{A}} = 1.33\,\text{k}\Omega \tag{12.14}$$

b)

$$R_{\text{eff}} = \frac{2\,\text{k}\Omega\,\mu\text{m}}{5\,\mu\text{m}} = 400\,\Omega \tag{12.15}$$

c)

$$R_{\text{eff}} = \frac{2\,\text{k}\Omega\,\mu\text{m}}{280\,\text{nm}} = \frac{2\,\text{k}\Omega\,\mu\text{m}}{0.28\,\mu\text{m}} = 7.1\,\text{k}\Omega \tag{12.16}$$

12.5 The CMOS inverter

Solution 5.1 *Problem is on page 13.*

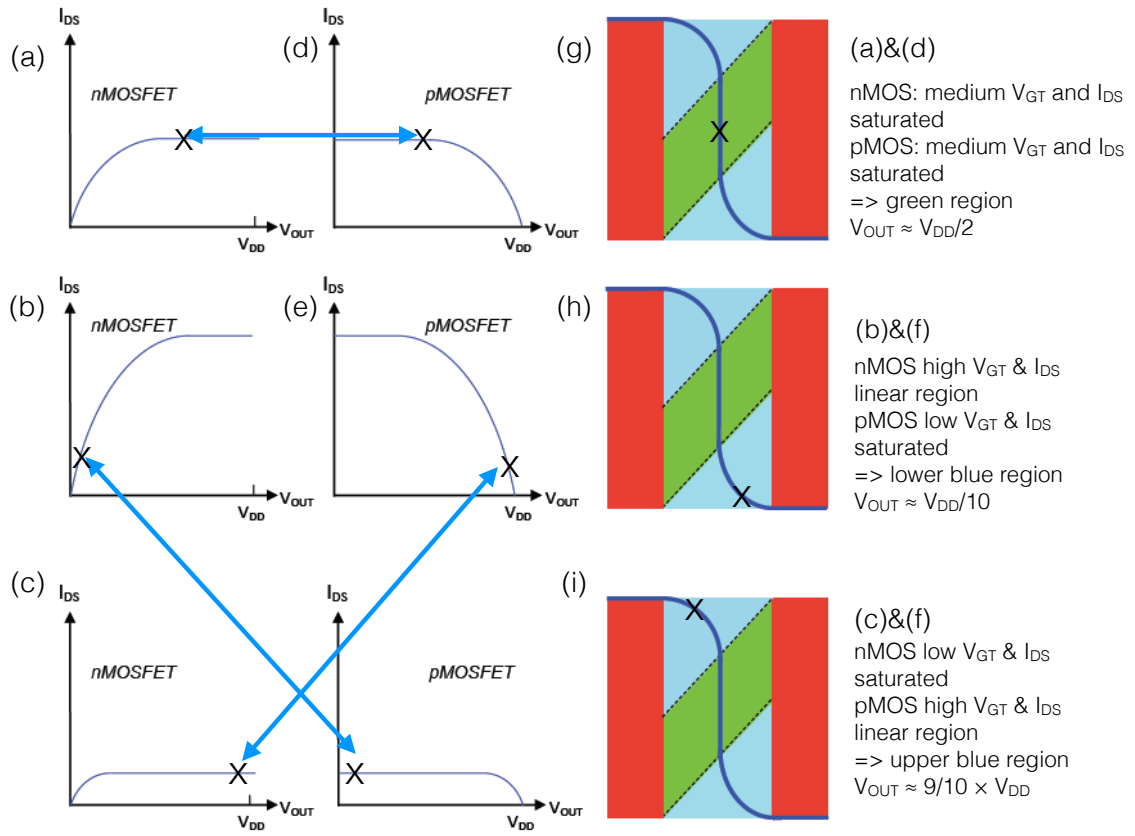


Figure 12.5: Solution to problem 5.1

.

The solution is shown in Figure 12.5.

Solution 5.2 *Problem is on page 13.*

a) In Figure 5.2 we see that there are three input voltages to the inverter where we have $V_{IN1} < V_{IN2} < V_{IN3}$. The lowest input voltage, V_{IN1}, corresponds to the lowest nMOS current, which is VGS1 in Figure 5.2. Similarly V_{IN2} corresponds to the VGS2 curve and V_{IN3} to the VGS3 curve. For the pMOS transistor we have the opposite situation: the lowest input voltage V_{IN2} corresponds to the highest VGS and thus the highest current etc.

b) To find the output voltage one has to consider the voltage at which the two current curves that correspond to the same input voltage cross. Thus, we find that for V_{IN1} the bias point is in region B, for V_{IN2} in region C, and for V_{IN3} in region D. See also Figure 12.5 where graph (i) corresponds to V_{IN1}, (g) to V_{IN2}, and (h) to V_{IN3}.

Solution 5.3 *Problem is on page 13.*

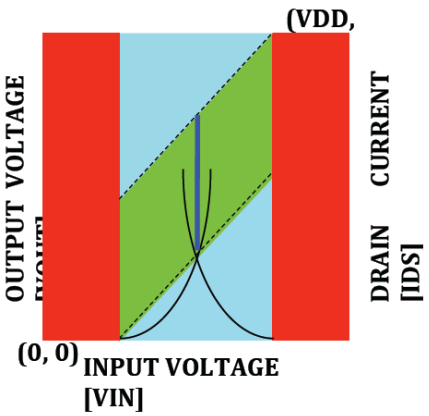


Figure 12.6: Solution to problem 5.3 a)

. .

a) In the red regions the current is zero (at least on the scale we draw it here). In the blue regions, the current equal to the saturation current for the transistor that is in the saturation region (that is the one with the lowest effective gate voltage, VGT. In the green region the current is equal to the saturation current when the two transistors have the same effective gate voltage. See Figure 12.6.

b) The transistor switching voltage, V_{sw}, is defined as the input voltage for which it is true that

$$V_{OUT} = V_{IN}. \tag{12.17}$$

That condition can only hold in the steep part of the VTC, that is the green area. The MOS current equation in saturation is:

$$I_{DS} = \frac{k}{2}(V_{GS} - V_T)^2 \tag{12.18}$$

This equation is the same for both nMOS and pMOS transistors, but with a minus sign for the pMOS transistor, which is because its current is due to holes rather than electrons. Kirchhoff's current law says that the sum of the currents (in or out) of the node is zero. In the output node we then get

$$I_{DSn} + I_{DSp} = 0. \tag{12.19}$$

Thus, we arrive at:

$$I_{DSn} = -I_{DSp} \tag{12.20}$$

which we rewrite as

$$\frac{k_n}{2}(V_{GSn} - V_{TP})^2 = \frac{k_p}{2}(V_{GSp} - V_{TP})^2. \tag{12.21}$$

To find V_{sw} we have to exchange V_G and V_S for the actual voltages in the inverter circuit. For both the nMOS and the pMOS transistor we have: $V_G = V_{sw}$. For the nMOS transtor we have $V_G = 0$ and for the pMOS

transistor $V_G = V_{\mathrm{DD}}$. Thus, we rewrite Eq. 12.21 as:

$$\frac{k_n}{2}(V_{\mathrm{sw}} - 0 - V_{TN})^2 = \frac{k_p}{2}(V_{\mathrm{sw}} - V_{\mathrm{DD}} - V_{TP})^2. \tag{12.22}$$

In this task we have the additional simplifying condition that $k \equiv k_n = k_p$ and $V_T \equiv V_{TN} = -V_{TP}$. Thus, we have:

$$\frac{k}{2}(V_{\mathrm{sw}} - V_T)^2 = \frac{k}{2}(V_{\mathrm{sw}} - V_{\mathrm{DD}} + V_T)^2, \tag{12.23}$$

or

$$(V_{\mathrm{sw}} - V_T)^2 = (V_{\mathrm{sw}} - V_{\mathrm{DD}} + V_T)^2. \tag{12.24}$$

In the next step one has to be careful, because the voltage that is squared for the nMOS transistor is positive while the voltage that is squared for the pMOS transistor is negative. So when we remove the squares we should keep the correct solutions of the two possibilities on each side. Finally we arrive at:

$$V_{\mathrm{sw}} - V_T = -(V_{\mathrm{sw}} - V_{\mathrm{DD}} + V_T), \tag{12.25}$$

which we rearrange to arrive at the solution:

$$V_{\mathrm{sw}} = \frac{V_{\mathrm{DD}}}{2} \tag{12.26}$$

c) Let us reason about it. If the nMOS transistor gives four times as much current for the same effective gate voltage (V_{GT}) then the point where the two transistors give the same current must happen at a lower input voltage than in b). Thus V_{sw} must be a bit lower than $\frac{V_{\mathrm{DD}}}{2}$. Because the currents are related to the square of the effective gate voltages the change in voltage will have to be related to the square root of the ratios between the current factors. We can write the current equation as:

$$\frac{k_n}{k_p}V_{GTn}^2 = V_{GTp}^2 \tag{12.27}$$

From this equation we see that with a ratio of four between the currect factors the ratio between the effective gate voltages has to be two. One way of seeing this is that of the available voltage range for effective gate voltages, $V_{\mathrm{DD}} - V_{TN} + V_{TP}$, one third is used by the nMOS transistor and two thirds by the pMOS transistor. Thus we have

$$V_{\mathrm{sw}} = V_{TN} + \frac{V_{\mathrm{DD}} - V_{TN} + V_{TP}}{3} = \frac{2V_{TN}}{3}V_{TN} + \frac{V_{\mathrm{DD}}}{3} + \frac{V_{TP}}{3} \tag{12.28}$$

A formal derivation is given under d) below.

d) To find V_{sw} in the general case we return to Eq. 12.22 which we can rearrange to:

$$\frac{k_n}{k_p}(V_{\mathrm{sw}} - V_{TN})^2 = (V_{\mathrm{sw}} - V_{\mathrm{DD}} - V_{TP})^2. \tag{12.29}$$

Again we are careful when removing the squares and thus arrive at:

$$\sqrt{\frac{k_n}{k_p}}(V_{\mathrm{sw}} - V_{TN}) = -(V_{\mathrm{sw}} - V_{\mathrm{DD}} - V_{TP}). \tag{12.30}$$

When we simplify we arrive at

$$V_{\mathrm{sw}} = \frac{V_{\mathrm{DD}} + V_{TP} + \sqrt{\frac{k_n}{k_p}}V_{TN}}{1 + \sqrt{\frac{k_n}{k_p}}} \tag{12.31}$$

Eq 12.31 is how the expression is most often written in textbooks. However, a slight rearrangement makes it much easier to remember and understand:

$$V_{\mathrm{sw}} = \frac{V_{\mathrm{DD}} + V_{TP} - V_{TN}}{1 + \sqrt{\frac{k_n}{k_p}}} + V_{TN} \tag{12.32}$$

From this formulation we see that our a bit informal reasoning in task c) has been verified more formally.

Solution 5.4 *Problem is on page 14.*

Solution will be added after prelab 1 is done.

Solution 5.5 *Problem is on page 15.*

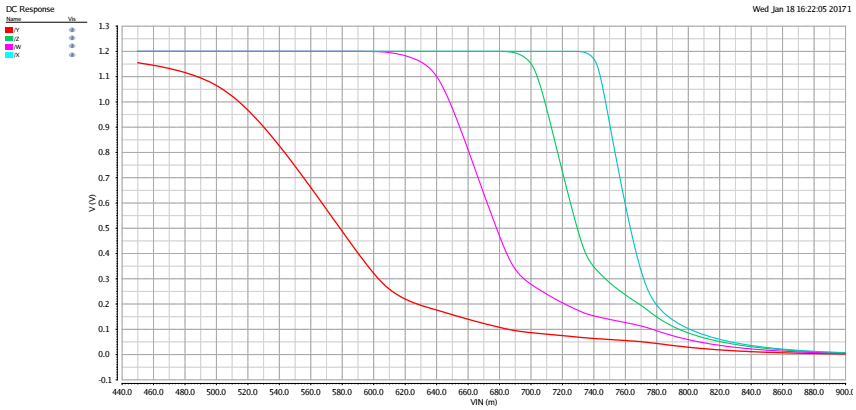


Figure 12.7: Plot of the four inverter curves resulting when all four inputs are connected together in the multi-input gate in Figure 12.3.

a) X (red) is Y4, Y(green) is Y3, and W(purple) is Y2. Motivation: With all inputs connected together we have four inverters with different p-to-n current and thus resistance ratios. Assuming that nMOS transistors have twice the current of pMOS transistors (which we know is very close to reality for the 65 nm process), the p-to-n resistance ratios for Y1–Y4 are $2R/R$, $R/2R$, $0.67R/3R$ and $0.5R/4R$. Since the threshold voltage are equal a 2-to-1 scaled inverter, that is one with R/R resistance ratio, should have $V_{sw} = V_{\mathrm{DD}}/2$. In Figure 5.5 we have $V_{sw} > V_{\mathrm{DD}}/2$ for all the curves. So then the VTC that is **not** drawn has to be the one where the p-to-n resistance ratio is larger than 1, that is curve Y1. The ordering of the three drawn VTCs has to be according to the p-to-n resistance ratios. In Figure 12.7 a diagram of all four VTCs is shown.

   Consequently, the VTC that is **not** drawn has to be the one where the p-to-n resistance ratio is larger than 1, that is Y1, which placed to the right of the other curves. The ordering of the three remaining VTCs has to be according to the p-to-n resistance ratios with Y4 having the lowest V_{sw} and then Y3 and Y2. In Figure 12.7 a diagram of all four VTCs is shown (unfortunately the line colors not the same as in the problem).

b) In Figure 12.8 below the two points that define the four voltages are shown. The resulting values are: For high level $MNH = V_{\mathrm{OHmin}} - V_{\mathrm{IHmin}} = 1.19\,\mathrm{V} - 0.832\,\mathrm{V} = 0.358\,\mathrm{V}$. For low level $MNL = V_{\mathrm{ILmax}} - V_{\mathrm{OLmax}} = 0.735\,\mathrm{V} - 0.05\,\mathrm{V} = 0.73\,\mathrm{V}$. As could be expected from the VTC the noise margins are not very equal so this inverter is not that well designed.

Solution 5.6 *Problem is on page 15.*

Solution 5.7 *Problem is on page 16.*

a) An ideal inverter has no parasitic capacitance at its output.

b) This delay is called the fanout-of-one delay, FO1. For an ideal inverter it is:

$$\mathrm{FO1} = 0.7 R_{\mathrm{eff}} C_{\mathrm{IN}} = 0.7 \frac{1.2\,\mathrm{V}}{500\,\mu\mathrm{A}/\mu\mathrm{m}} \times 3 \times 1.3\,\mathrm{fF}/\mu\mathrm{m} = 6.5\,\mathrm{ps}. \tag{12.33}$$

   The factor 3 for the capacitance in the expression is due to the inverter input capacitance being three times as large as the gate capacitance for that of only the nMOS transistor. Note that the delay is the same regardless of the size of the two inverters.

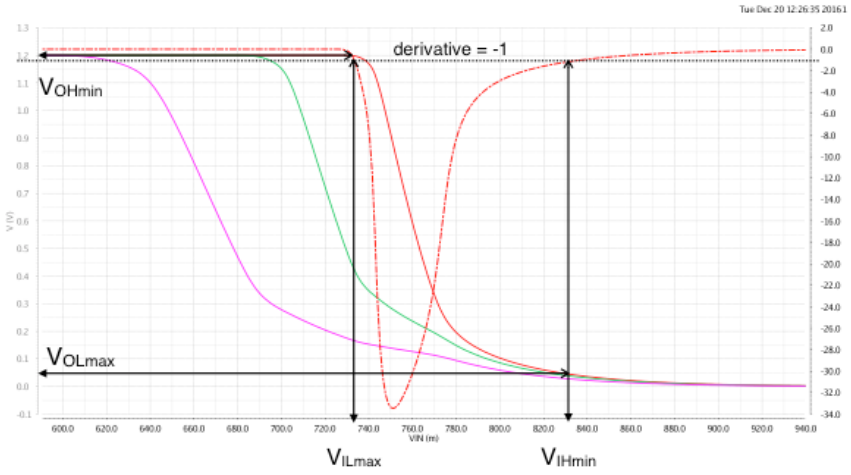


Figure 12.8: VTC with voltages necessary to calculated noise margins for output X indicated.

Solution 5.8 *Problem is on page 16.*

a) The fanout-of-four delay is given by this equation:

$$\mathrm{FO4} = 0.7RC(4 + p). \tag{12.34}$$

In the 0.35 µm process we find:

$$0.7RC = 0.7 \times 6\,\mathrm{k\Omega\mu m} \times 6\,\mathrm{fF/\mu m} = 25\,\mathrm{ps}. \tag{12.35}$$

With $p = 1$ we thus arrive at $\mathrm{FO4} = 125\,\mathrm{ps}$.

b)

Solution 5.9 *Problem is on page 17.*

a) Check which inverter we refer to here.

b) The resulting delay equation is then

$$t_{pd} = R'C. \tag{12.36}$$

Comment: When you only consider gate delays this simplification seems like a good idea (and it is the way is done in Weste & Harris) because it simplifies the delay equations. However, when we introduce wires with real physical resistances, it gets messy if one treats gates and wires differently. So therefore we do not do so in this course. We have to carry the 0.7 factor with us in delay calculations.

Solution 5.10 *Problem is on page 17.*

a)

b)

c)

d)

e)

Solution 5.11 *Problem is on page 17.*

a)

b)

c)

Solution 5.12 *Problem is on page 18.*

Solution 5.13 *Problem is on page 18.*

a)

b)

Solution 5.14 *Problem is on page 18.*

a) We know that for minimum delay each stage should have the same effort and that a stage effort, f, of 4 is good to minimize the delay. 1024 is 2^{10}, which is also $4 \times 4 \times 4 \times 4 \times 4$, so 5 inverters in all and thus four inverters in the box is a good solution.

b) The delay is $p + gh$ or $p + f$. We have $f = 4$ (see a) above) and $p_{\text{inv}} = 0.5$ (from problem statement). So the normalized delay, d, is $5 \times (4 + 0.5) = 22.5$ and with $\tau = 4$ ps we have a delay of $d \times \tau = 22.5 * 4 = 90$ ps.

e) BONUS QUESTION It would be better to remove one inverter than to add one inverter since the dynamic power will be lower while the delay would only be negligibly longer.

12.6 Delay for complex gates and paths

Solution 6.1 *Problem is on page 19.*

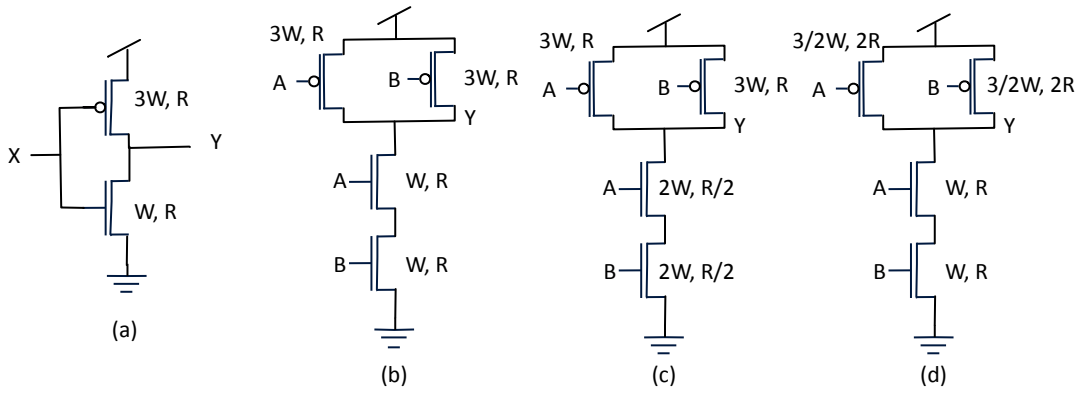


Figure 12.9: (a) Inverter with 3-to-1 scaling for same resistance, (b) NAND2 gate with same scaling as inverter, (c) and (d), two scalings that make the worst-case nMOS and pMOS resistances the same.

a) In this unusual CMOS process the pMOS transistor in the inverter has to be three times as wide as the nMOS transistor to give the same saturation current, that is to have the same effective resistance. Such an inverter is shown in Figure 12.9(a). Assuming that the oxide capacitance, C_{ox}, is the same for nMOS and pMOS transistors (which we always assume), that means the when the transistor is scaled electrically symmetrically the pMOS transistor accounts for three fourths of the inverter input capacitance and the nMOS transistor for one fourth. It is then convenient to use:

$$C_{inINV} = 3C + 1C = 4C \tag{12.37}$$

where C is the capacitance for any type transistor with width W. The RC product for the inverter can then be written as

$$RC_{INV} = R \times C_{inINV} = R \times 4C \tag{12.38}$$

Note, that we use $4C$ just for our convenience. It is not necessary to do so.

The NAND2 gate is shown in Figure 12.9(b) with the same transistor widths and thus resistances as for the inverter in Figure 12.9(a). For the NAND2 gate in Figure 12.9(b) the worst-case resistance (that is the highest resitance) for the n-net is $2R$ while for the p-net it is R. Therefore, we must scale the nMOS transistors relative to the pMOS transistors.Figures 12.9(c) and (d) shows two different ways of scaling. In Figure 12.9(c) the nMOS transistor are made wider so the worst-case resistance is R also in the n-net. In Figure 12.9(d) the pMOS transistors were instead made narrower which makes the worst-case resistance of the p-net also $2R$. From Figure 12.9(c) we find

$$RC_{NAND2} = R_{NAND2} \times C_{inNAND2} = R \times (2C + 3C) = 5RC \tag{12.39}$$

From Figure 12.9(d) we find

$$RC_{NAND2} = R_{NAND2} \times C_{inNAND2} = 2R \times (C + \frac{3}{2}C) = 5RC \tag{12.40}$$

So we see that the two scalings give the same RC product for the NAND2 gate. The logical effort is defined as:

$$g = \frac{R_{gate} \times C_{gateinput}}{R_{inv} \times C_{inv}} \tag{12.41}$$

In this case we get $g_{NAND2} = 5/4$, which holds for both input A and B since they have the same transistor widths and thus the same gate capacitance. If the resitances are the same in the inverter and in the complex gate Equation simplifies to:

$$g = \frac{C_{gateinput}}{C_{inv}} \tag{12.42}$$

The parasitic delay, p, is the part of the delay that does not depend on the load capacitance, that is the constant, or internal part. For a static CMOS gate that is the part of the delay that is due to the parasitic capacitances at the drains of the output transistors. The definition for p is:

$$p = \frac{R_{gate} \times C_{parasitic}}{R_{inv} \times C_{inv}} \tag{12.43}$$

In this case the problems states that $p_{\text{inv}} = 0.5$, which means that the transistor output capacitances are halv of the intput capacitances. As for the electrical effort the equation can be simplified if the resistances are the same. Thus, we arrive at:

$$p = \frac{C_{parasitic}}{C_{inv}} = \frac{p_{\text{inv}} C_{\text{gate -to-output}}}{C_{inv}}, \tag{12.44}$$

where $C_{\text{gate-to-output}}$ is the gate capacitance of all transistors connected to the output node. For the NAND2 gate in this strange process we find

$$p_{NAND2} = \frac{0.5 \times (3 + 3 + 2)C}{4C} = 1 \tag{12.45}$$

For the NOR2 gate we just give the answers:

$$g_{NOR2} = \frac{C + 6C}{4C} = \frac{7}{4} \tag{12.46}$$

$$p_{NOR2} = \frac{0.5 \times (6 + 1 + 1)C}{4C} = 1 \tag{12.47}$$

b) In task a) we found that the NAND2 gate in this strange process has $g_{NAND2} = 5/4$ and $p_{NAND2} = 1$. The normalized delay is defined as:

$$d = g \times h + p, \tag{12.48}$$

where h is the electrical effort which can only be known when a load is connected to the gate. The electrical effort is defined as:

$$h = \frac{C_{\text{load}}}{C_{\text{in}}} \tag{12.49}$$

With the scaling stated in the problem, we have the inverter input capacitance, which is the capacitance loading the NAND2 gate, as $6C + 2C = 8C$, whereas the input capacitance for both inputs of the NAND2 case is $3C + 2C = 5C$. Consequently the normalized delay for the NAND2 gate with these transistorsizes is

$$d_{NAND2} = \frac{5}{4} \times \frac{8C}{5C} + 1 = 3 \tag{12.50}$$

Note that the normalized delay is expressed in units of τ, so one has to multiply with τ to find the delay in seconds. For this process we do not know the value of τ.

Solution 6.2 *Problem is on page 19.*

We use the scaling that gives the same effective resistance as for the inverter wich has width 2 for the pMOS transistor and 1 for the nMOS transistors.

a) In this gate all nMOS transistors have width 2. The pMOS transistors connected to A, B and C have widths 6, whereas the one connected to D has width 2. Then we find, for A, B and C inputs:

$$g_{A,B,C} = \frac{6C + 2C}{3C} = \frac{8}{3} \tag{12.51}$$

For the D input:

$$g_D = \frac{2C + 2C}{3C} = \frac{4}{3} \tag{12.52}$$

In the problem it was not specified what p_{inv} is, so we express the parasitic delay in pinv:

$$p = p_{\text{inv}} \frac{(6 + 2 + 2)C}{3C} = \frac{10}{3} p_{\text{inv}} \tag{12.53}$$

b) For the second gate all pMOS transistors have width 4, the two nMOS transistors in series to the left have width 2 and all other nMOS transistors have widths 3.

$$g_{A_2,B_2} = \frac{2C + 3C + 4C + 4C}{3C} = \frac{13}{3} \tag{12.54}$$

For the D input:

$$g_{A_1,B_1} = \frac{4C + 3C}{3C} = \frac{7}{3} \tag{12.55}$$

In the problem it was not specified what p_{inv} is, so we express the parasitic delay in pinv:

$$p = p_{\text{inv}} \frac{(4 + 4 + 4 + 2 + 3)C}{3C} = \frac{17}{3} p_{\text{inv}} \tag{12.56}$$

Solution 6.3 *Problem is on page 19.*

There are many possibilities. The requirement stated in the problem is that the worst-case n-net resistance and p-net resistance should be the same for each of the four gates, but there is no requirement that they should all be the same (that is probably impossible). One solution is to make all the nMOS transistors the same width; let's call it W and then scale the pMOS transistors to match. We assume that such an nMOS transistor has the resistance R. Then the n-net worst-case resistance is R for output Y_1 (inverter); 2R for output Y_2 (NAND2), 3R for output Y_3 (NAND3), and 4R for output Y_4 (NAND3). Assuming that the pMOS transistors are half as strong as the nMOS transistors we then must use the widths 2W for transistor 161, W for transistors numbered 15X, 2W/3 for pMOS transistors numbered 14X and W/2 for pMOS transistors numbered 13X.

Solution 6.4 *Problem is on page 19.*

For the NAND3 gate the logical effort, g_{NAND3}, is 5/3 for all three inputs. With $p_{\text{inv}} = 1$, the parasitic delay, p_{NAND3}, is 3.

The normalized delay for one stage can be written as:

$$d = g \times h + p. \tag{12.57}$$

From Figure 6.2 we see that the electrical effort with branching for the 2-input NAND gates is:

$$h_{NAND2} = \frac{3 \times C_{inNAND3}}{C_{inNAND2}} = \frac{3 \times 8C}{8C} = 3. \tag{12.58}$$

And for the 3-input NAND gates the electrical effort is:

$$h_{NAND3} = \frac{2 \times C_{inNOR2}}{C_{inNAND3}} = \frac{2 \times 16C}{8C} = 4. \tag{12.59}$$

And for the last stages, the NOR2 gates, we have:

$$h_{NOR2} = \frac{C_{\text{LOAD}}}{C_{inNOR2}} = \frac{48C}{16C} = 3. \tag{12.60}$$

To find the total normalized delay we can just sum up the normalized delay for each stage, which we now number, 1, 2, 3 from the input for simplicity:

$$d_{\text{tot}} = d_1 + d_2 + d_3 = \frac{4}{3} \times 3 + 2 + \frac{5}{3} \times 4 + 3 + \frac{5}{3} \times 3 + 2 = 22\frac{2}{3} \tag{12.61}$$

Solution 6.5 *Problem is on page 20.*

We already have the logical efforts and the parasitic delays for all gates from the solution to the previous problem. To find the best scaling we need to calculate the path effort:

$$F = G \times H \times B. \tag{12.62}$$

We have the path logical effort:

$$G = g_1 \times g_2 \times g_3 = \frac{4}{3} \times \frac{5}{3} \times \frac{5}{3}, \tag{12.63}$$

the path electrical effort:

$$H = \frac{C_{\text{out-for-path}}}{C_{\text{in-for-path}}} = \frac{45C}{8C} = \frac{45}{8}, \tag{12.64}$$

and finally the path branching effort:

$$B = b_1 \times b_2, \tag{12.65}$$

where branching effort for each stage is defined as:

$$b_i = \frac{C_{\text{onpath}} + C_{\text{offpath}}}{C_{\text{onpath}}}. \tag{12.66}$$

Note that one usually does not have a branching effort for the last stage since all capacitance in on the path for that stage. So in this problem we have

$$b_1 = \frac{8C + 2 \times 8C}{8C} = 3, \tag{12.67}$$

and

$$b_2 = \frac{16C + 16C}{16C} = 2. \tag{12.68}$$

Thus, we arrive at

$$B = 3 \times 2 = 6. \tag{12.69}$$

All in all we find

$$F = G \times H \times B = \frac{4}{3} \times \frac{5}{3} \times \frac{5}{3} \times \frac{45}{8} \times 6 = 5^3 \tag{12.70}$$

We know that the optimum is when the stage effort in each stage is the same, and here it is obvious that $f_{\text{opt}} = 5$ since we have three stages. Once we have f_{opt} we can immediately find the optimum delay as

$$D = N \times f_{\text{opt}} + \sum p_i \tag{12.71}$$

which in this problem is:

$$D = 3 \times 5 + 7 = 22 \tag{12.72}$$

The sizes (input capacitances) for the three stages can be found starting either from the input or from the output of the path. Here we choose the output. For the third stage we have with $f_{\text{opt}} = 5$

$$5 = \frac{5}{3} \times \frac{45C}{C_{\text{in3}}}, \tag{12.73}$$

which results in $C_{in3} = 15C$. Similarly for the second stage we have:

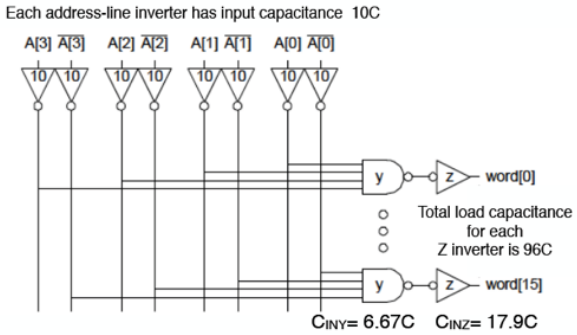
$$5 = \frac{5}{3} \times \frac{2 \times 15C}{C_{in2}}, \tag{12.74}$$

which gives us $C_{in2} = 10C$. The first stage in a path is not scaled; however, it is good practice to check that the effort for that stage also becomes f_{opt} when we put in the numbers. For our first stage we find:

$$f_1 = \frac{4}{3} \times \frac{3 \times 10C}{8C} = 5. \tag{12.75}$$

Solution 6.6 *Problem is on page 20.*

a) Below is the detailed schematic of the decoder circuitry with the resulting sizes for task a).



This is a problem where path delay is applicable. The driver inverters each drive eight wires connected to eight of the nand gates (to half of the 16 words). The load capacitance for the z inverter is $32 \times 3C = 96C$. The logical effort for the 4-input nand gate is 2 (which one may have to arrive at from the circuit schematics, though not included here). So for the path we have $G = 1 \times 2 \times 1$, $B = 8 \times 1 \times 1$ and $H = \frac{96C}{10C} = 9.6$. All in all the path effort, $F = GBH$ is $2 \times 8 \times 9.6 = 153.6$. The optimal stage effort is then $f_{opt} = \sqrt[3]{153.6} = 5.36$. We can find the sizing by starting from the output or the input, but usually it is easier to start from the output. We use the relation $f_{opt} = g\,h$ for each gate and determine the input capacitance that makes this relation true. Inverter z should accordingly have an input capacitance $C_{inz} = \frac{96C}{5.36} = 17.9C$. The 4-input nand gate should have an input capacitance of $C_{iny} = \frac{2 \cdot 17.9C}{5.36} = 6.67C$.

We should also check that the relationship holds for the first inverter: $C_{inx} = \frac{8 \cdot 6.67C}{5.36} = 9.95C$. It is not exactly $10C$ because there has been some rounding in the calculations, but it is close enough to convince us that we did not make any calculation mistake.

b) The normalized delay $D = 3\,f_{opt} + \sum p$. The inverter has a parasitic delay of $p_{inv} = 1$. But the part we do not know is the parasitic delay of the 4-input nand gate. From the schematic we deduce that the parasitic delay for a scaled 4-input nand gate is 4 (this calculation is not included here). Thus we arrive at $D = 3 \times 5.36 + 1 + 4 + 1 = 22.08$. In the 65 nm process which has $\tau = 5$ ps we would thus have a delay of 110 ps.

c) The difference from the case in task a) is that the word capacitance is doubled which doubles H. In this case we have $F = GBH$, $2 \times 8 \times 19.2 = 307.2$. With four stages we get $f_{opt} = \sqrt[4]{307.2} = 4.18$. The resulting delay with parasitics is then $= 4 \times 4.18 + 1 + 4 + 1 + 1 = 23.72$. With three stages we instead have $f_{opt} = \sqrt[3]{307.2} = 6.74$ and a total delay of $3 \times 6.74 + 1 + 4 + 1 = 26.22$. So the answer is yes, the delay will be shorter with an extra inverter.

12.7 Wire delay

Solution 7.1 *Problem is on page 23.*

a) Sum the currents into (or out of) each circuit node to find the nodal equations:

$$0 = \frac{v_1(t) - V_S}{R_1} + \frac{v_1(t) - v_2(t)}{R_2} + C_1 \frac{dv_1(t)}{dt} \tag{12.76}$$

$$0 = \frac{v_2(t) - v_1(t)}{R_2} + C_2 \frac{dv_2(t)}{dt} \tag{12.77}$$

where V_S is the source voltage.

b)

c)

Solution 7.2 *Problem is on page 23.*

a) The wire capacitance is:

$$C_w = 0.2\,\mu\text{m} \times 25\,\mu\text{m} \times 0.4\,\text{fF}/\mu\text{m}^2 = 2\,\text{fF}. \tag{12.78}$$

The wire resistance is:

$$R_w = \frac{25\,\mu\text{m}}{0.2\,\mu\text{m}} \times 0.2\,\Omega/\square = 25\,\Omega. \tag{12.79}$$

b) For the wire that has a width of 100 nm we have

$$c = 0.1\,\mu\text{m} \times 0.4\,\text{fF}/\mu\text{m}^2 = 0.04\,\text{fF}/\mu\text{m}, \tag{12.80}$$

and

$$r = \frac{1\,\mu\text{m}}{0.1\,\mu\text{m}} \times 0.2\,\Omega/\square = 2\,\Omega/\mu\text{m}. \tag{12.81}$$

The critical length for wire insertion with $p_{inv} = 1$ is:

$$L_{\text{crit}} = 2\sqrt{\frac{RC}{rc}} \tag{12.82}$$

which we can also write as

$$L_{\text{crit}} = 2\sqrt{\frac{t_{\text{rep}}}{rc}} \tag{12.83}$$

Thus, we find:

$$L_{\text{crit}} = 2\sqrt{\frac{4.6\,\text{ps}}{0.08\,\text{fs}/\mu\text{m}^2}} = 2\sqrt{\frac{4600\,\text{fs}}{0.08\,\text{fs}/\mu\text{m}^2}} = 480\,\mu\text{m} \tag{12.84}$$

Solution 7.3 *Problem is on page 23.*

a) Answer: $R = 1\,\text{k}\Omega, C = 650\,\text{fF}$, calculated from $r = 100\,\Omega/\text{mm}, C = 650\,\text{fF}/\text{mm}$.

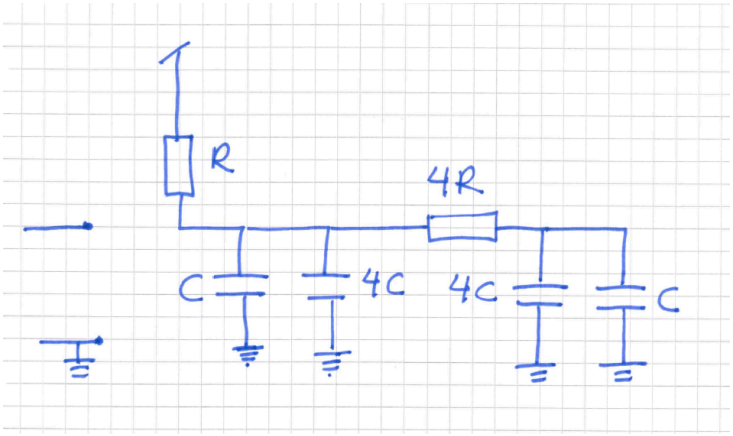
b) Answer:

$$2\,\text{k}\Omega \times (3.25 + 650 + 3.25)\text{fF} + 1\,\text{k}\Omega \times (325 + 3.25)\text{fF} = 1.64\,\text{ns} \tag{12.85}$$

c) Answer: $R_{\text{rep}}C_{\text{rep}} = 6.5\,\text{ps}$, $rc = 6.5\,\text{ps}/\text{mm}^2$. Together they give $L_{\text{crit}} = 2\sqrt{\frac{R_{\text{rep}}C_{\text{rep}}}{rc}} = 2\sqrt{\frac{6.5}{6.5}} = 2\,\text{mm}$.

Solution 7.4 *Problem is on page 24.*

a) The sketch is shown below:



Note that the input of the inverter is not connected to the inverter output. Here we have not even drawn in the capacitance at inverter input since it does not influnce the dealy at all.

b) The RC product is:

$$\tau = R \times 10C + 4R \times 5C = 30RC \tag{12.86}$$

Thus, the propagation delay is $0.7 \times 30RC$.

c) The higher receiver input capacitance results in a RC prodcruct of:

$$\tau = R \times 11C + 4R \times 6C = 35RC \tag{12.87}$$

Thus, there is an *increase* in the propagation delay of $0.7 \times 5RC$.

d) A 2-input NAND gate has $g = 4/3$ and $p = 2$.That means that its parasitic capacitance is 3/2 times its input capacitance (if p_{inv} is 1 which is the case in this problem). Thus, in this case the parasitic capacitance of the NAND gate is 3C. When we are driving wires we also need to calculate the resistance of the gate explicitly. We know from the definition that $g = \frac{RC_{\text{gate}}}{RC_{\text{inv}}}$. In this case we thus find:

$$\frac{R_{NAND2}}{R_{inv}} = g \times \frac{C_{inv}}{C_{NAND2}} = 4/3 \times 1/2 = 2/3. \tag{12.88}$$

Thus, the new RC product is:

$$\frac{2}{3}R \times 12C + 4R \times 5C = 26RC \tag{12.89}$$

In this case there is a *decrease* in the propagation delay of $0.7 \times 4RC$.s

e) In the Elmore delay model a branch contributes its entire capacitance multiplied with the resistance up to the point where the branch connects to the main path. In this case the added capacitance is $2C$ and the resistance to the midpoint is $R + 2R$. Thus, the result of adding the branch is an increase in the propagation delay of $0.7 \times 6RC$.

f) We denote the RC product for the inverter with t_{inv} since it is a constant. Then we have $R = \frac{t_{\text{inv}}}{xC}$ and $C = \frac{xt_{\text{inv}}}{R}$ where x is the size of the inverter. The general RC expression for the delay is then

$$\tau = \frac{t_{\text{inv}}}{xC}\left(\frac{2xt_{\text{inv}}}{R} + C_w\right) + R_w\left(\frac{xt_{\text{inv}}}{R} + \frac{C_w}{2}\right), \tag{12.90}$$

which we can simplify to

$$\tau = \frac{2t_{\text{inv}}^2}{RC} + \frac{t_{\text{inv}}C_w}{xC} + \frac{xt_{\text{inv}}R_w}{R} + \frac{R_wC_w}{2}. \tag{12.91}$$

Only the two middle terms depend on x. We find the minimum for

$$x_{\text{opt}} = \sqrt{\frac{RC_w}{R_wC}} \tag{12.92}$$

Note: In in this kind of problem in an exam it would not be required to derive the expression for the optimum size if you remember it, but it may actually be easier than to memorize it.

In this particular case we have

$$x_{\text{opt}} = \sqrt{\frac{R \times 8C}{4R \times C}} = \sqrt{2} \tag{12.93}$$

So the size of the inverter should be such that its resistance is $\frac{R}{\sqrt{2}}$ and its capacitance is $\sqrt{2}C$. The RC product for the wire and driver is then,

$$\tau = 2RC + 4\sqrt{2}RC + 4\sqrt{2}RC + 16RC \approx 29.3RC, \tag{12.94}$$

and the propagation delay is $0.7 \times 29.3RC$.

So we did not gain much by optimizing the original setup.

Solution 7.5 *Problem is on page 24.*

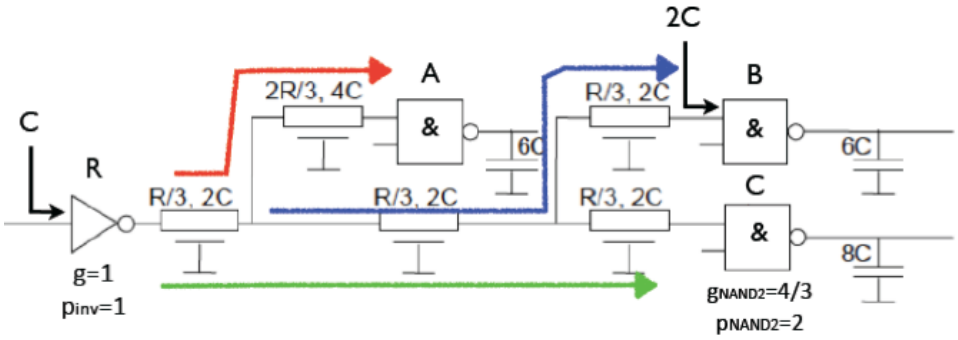


Figure 12.10: Clock-distribution network with a driver driving three receivers over different wires with the three paths indicated. The input capacitance of the NAND gates is 2C.

a) In this circuit there are three paths to the clock gaters A, B and C (red, blue and green in Figure 12.10).

We need to model the wire segments in these paths using the pi model and calculate the wire delay using Elmore's model. However, the clock skew is the **difference** in delay. Those parts of the delays that are the same for all three paths we do not have to calculate. The Elmore delay can be calculated as the delay due to the main path plus the delay due to branches. In this circuit the main path due to the wire segments to the inputs A, B, and C all have R and 6C and the input capacitances are also all the same: 2C. Thus, we do not have to calculate the main-path delay. However, the branch delays differ. There are three branches and each path incorporates two of them. Remember that in Elmore's model the entire capacitance of a branch should be multiplied with the resistance up to the point where the branch starts.

Branch RC delay due to red branch: $bd_A = \left(R + \frac{R}{3}\right) \times 6C = \frac{24}{3}RC$

Branch RC delay due to blue, green branches: $bd_B = bd_C = \left(R + \frac{2R}{3}\right) \times 4C = \frac{20}{3}RC$

Total branch RC delay to node A is: $bd_B + bd_C = 2 \times \frac{20}{3}RC = \frac{40}{3}RC$

Total branch delay to nodes B and C are the same: $\left(\frac{24}{3} + \frac{20}{3}\right)RC = \frac{44}{3}RC$

So the **clock skew is** $0.7 \times \frac{4}{3}RC$ and the delay to clock gaters B and C are the longest ones, whereas the delay to gater A is shorter.

b) The difference in load capacitance between clock gaters B and C is what will cause clock skew between the outputs of gaters B and C since they are scaled the same. That difference is 2C. The delay to the gate inputs will not change.

Since there is no wires to account for at the outputs of the NAND gates the delay can be calculated using logical effort:

$$\tau_{skew} = g_{\text{NAND2}} \times (h_{\text{C}} - h_{\text{B}}) \tag{12.95}$$

where we were careful to take the difference in the order that makes the skew positive. Equation 12.95 evalutes to 4/3 because we have:

$$h_C - h_B = \frac{8C}{2C} - \frac{6C}{2C} = \frac{2C}{2C} = 1. \tag{12.96}$$

We can also calculate R explicitly which is done below.

Gaters A, B and C are identical. The logical effort, g, is defined as "*the ratio of the input capacitance to that of an inverter that can drive the same current* " or, in other words, the same effective resistance. That is, gC corresponds to R. Here we have that $C_{NAND} = 2C = g\frac{3}{2}C$ and thus $R_{NAND} = \frac{2}{3}R$. So the resulting clock skew is $0.7 \times \frac{2}{3}R \times 2C$= **is** $0.7 \times \frac{4}{3}RC$.

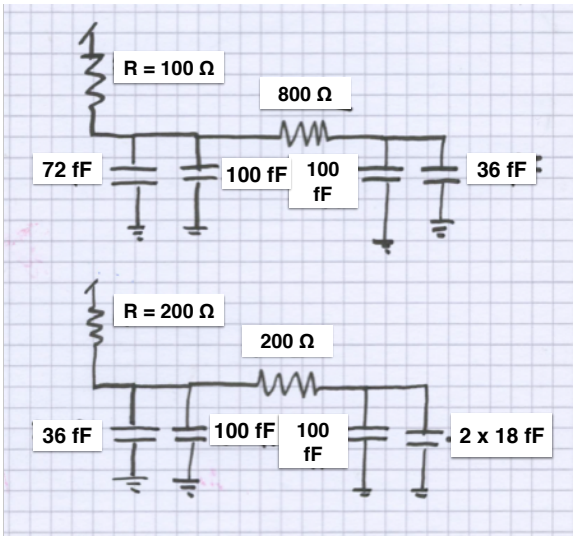
ADDITION: We did not have to calculate the main path delay, but it could be of interest to do so anyway for practice. To do so we need only one wire segment with resistance R and capacitance 6C. With the usual model for a driver that drives a a receiver over a wire we fine the RC product to be:

$$R \times (C + 3C) + R \times (3C + 2C) = 9RC \tag{12.97}$$

Thus, the resulting delay for the main path is $t_{pd} = 0.7 \times 9RC$.

Solution 7.6 *Problem is on page 24.*

There are two parts in this circuit as is shown in the figure below:



For the second part we use a collapsed tree in this solution, but one can also use Elmore branches.

a) Elmore for first part: $100\,\Omega \times (72 + 100 + 100 + 36)$fF + $800\,\Omega \times (100 + 36)$ fF = 139.6 ps. Elmore for the second part: $200\,\Omega \times (36 + 100 + 100 + 36)$ fF + $200\,\Omega \times (100 + 36)$fF = $3 \times 200 \times 136$fF = 81.6 ps. All in all the delay becomes $td = 0.7 \times (139.6 + 81.6) =$ **155 ps** .

b) We know that for the 2-input NAND gate we have that g_{NAND2} is 4/3 and p_{NAND2} is 2 (otherwise we could derive these numbers). The path effort from A to B is then $F = G \times H = 4/3 \times 72 \times 1.5 = 64$. For minimum delay all stage efforts should be the same. In this case all stage efforts, f, should be 4 since $4 \times 4 \times 4 = 64$. To find the inverter sizes we can start from the output or the input of the path. From the input we have $4/3 \times h_{NAND} = 4 \Rightarrow h_{NAND} = 3$ so the input capacitance of the first inverter in the buffer should be **4.5 fF**. The input of the second inverter should be 4×4.5fF = **18 fF**. These capacitances correspond to drive strenghs 200 / 4 = **50X** for the second inverter and 50/4 = **12.5X** for the first inverter.

c) The resulting normalized delay, d, is 4+4+4+P where P is the sum of the parasitic delays for the three gates. We have $P = 2 + 1 + 1 = 4$. Here we use our prior knowledge that the 2-input NAND gate has $p = 2$, but we could also derive it, if we did not remember. Thus, we have $d = 16$. To calculate the delay in seconds we also need τ. But what is τ in this process? We had better check that too. It is $0.7\,RC = 0.7 \times 72$ fF$\times 0.1$ kΩ = 5 ps. So, the propagation delay from A to B is 16×5 ps = **80 ps**.

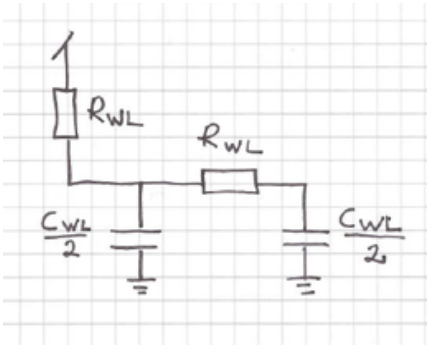
COMMENT: Is the total delay from A to C now 80 ps + 155 ps = 235 ps or is there some delay we have not accounted for, since we assumed infinite drive strength at point B in task a)? In our path delay calculation in b) we accounted for the electrical effort of the second buffer so we have accounted for that delay and therefore it is correct to assume that the total delay is the sum of the two delays.

Solution 7.7 *Problem is on page 25.*

a) The number of squares for one WL wire is its length/width = 128 / 0.1 = 1280. Thus, the resistance of the WL is $R_{\mathrm{WL}} = 1280 \times 0.1\ \Omega/\square = 128\,\Omega$.

b) The capacitance of one WL is $C_{\mathrm{WL}} =$ length $\times (C_{\mathrm{GND}} + 2 \times C_{\mathrm{INTERWIRE}}) + \#cells \times 2 \times C_{\mathrm{G}}$ fF. In this case we have $C_{\mathrm{WL}} = 128\,\mu\mathrm{m} \times (0.1\,\mathrm{fF}/\mu\mathrm{m} + 2 \times 0.02\,\mathrm{fF}/\mu\mathrm{m}) + 128 \times (2 \times 0.1\,\mathrm{fF}) = 128 \times 0.34\,\mathrm{fF} = 43.5\,\mathrm{fF}$.

c) Here is the model:



The delay can be computed as $t_{dWL} = 3/2 \times R_{WL}C_{WL} = 3.7\,\mathrm{ps}$.

d) The energy is computed as $E_{\mathrm{WL}} = C_{\mathrm{WL}}V_{\mathrm{DD}}^2 = 43.5\,\mathrm{pJ}$ since only one WL is charged for each reading of the memory.

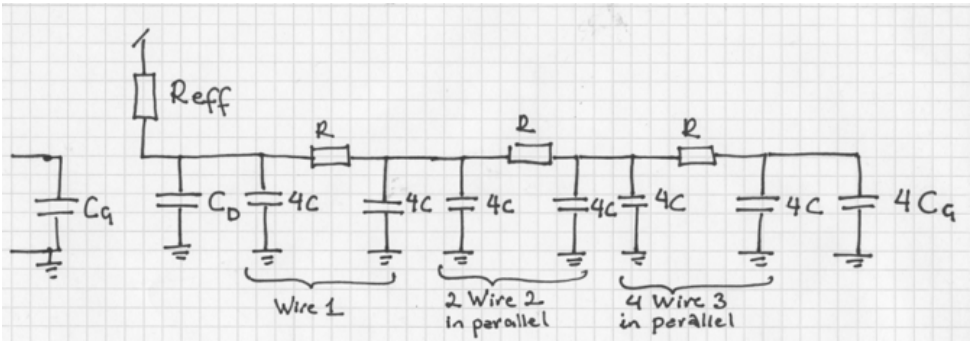
e) Resistance: The length of the wire is halved, but its width is not changed since the wire already has the minimum width. Thus, we have half the number of squares which gives $R_{\mathrm{WL2}} = R_{\mathrm{WL}}/2 = 64\,\Omega$.

Capacitance: The parallel M2 wires are now approximately at half the distance they were before. If we assume plate capacitances, the capacitance doubles so we have $C_{\mathrm{INTERWIRE2}} = 2 \times C_{\mathrm{INTERWIRE}}$. C_{GND} and C_{G}, and the number of cells remain the same. So in this case we get: $C_{\mathrm{WL2}} = 64\,\mu\mathrm{m} \times (0.1\,\mathrm{fF}/\mu\mathrm{m} + 2 \times 0.04\,\mathrm{fF}/\mu\mathrm{m}) + 128 \times (2 \times 0.1\,\mathrm{fF}) = 64 \times 0.58\,\mathrm{fF} = 37.1\,\mathrm{fF}$ The shorter WL wires improved the resistance much more than the capacitance.

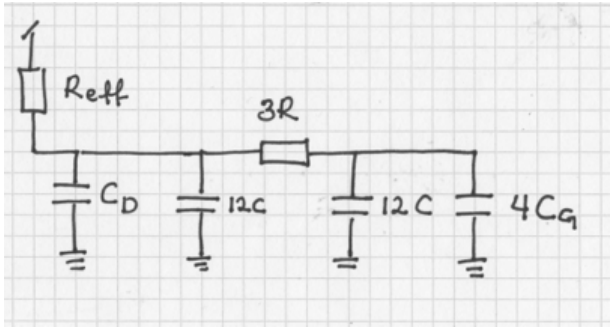
The new delay is $t_{\mathrm{dWL2}} = 3/2 \times 2.37 = 1.6\,\mathrm{ps}$. The new energy E_{WL2} is 37 pJ. So making the memory smaller more than halved the delay, but the energy required is almost the same. This result is due to the fact that the main improvement was the halved resistance whereas the capacitance remained almost the same.

Solution 7.8 *Problem is on page 26.*

This problem can be solved by collapsing the H-tree; it is also possible apply Elmore's formula directly and handle the branches explicitly. Here we use the collapsed tree. If you have not yet done lab 4 consult the instructions for prelab 4 to read more about how to collapse the tree. The resulting schematic is shown in the figure below:



Since the three resulting wire segments are identical we can merge them into one segment to further simplify the calculations (You may remember from lab 4, that in the Cadence simulations the number of segments a wire is divided into makes a difference in the delay calculations, but in the hand calculations it does not). The further simplified schematic is shown below:



a) We use Elmore's formula to find this expression for t'_{dFO4} (the delay without the 0.7 factor):

$$t'_{\text{dFO4}} = R_{\text{eff}}\, C_{\text{D}} + R_{\text{eff}}\, 24C + R_{\text{eff}}\, 4C_{\text{G}} + 3R\, 12C + 3R\, 4C_{\text{G}} \tag{12.98}$$

This expression we can simplify further using $C_{\text{D}} = C_{\text{G}}$ and rearranging:

$$t'_{\text{dFO4}} = 5R_{\text{eff}}C_G + 24\, R_{\text{eff}}\, C + 36RC + 12RC_{\text{G}} \tag{12.99}$$

We have that the FO4 delay is $0.7t'_{\text{dFO4}}$. We identify $\tau' = R_{\text{eff}}C_{\text{G}}$, which we know is a process constant; we then see that the first term in corresponds to the usual FO4 delay.

b) To find the optimal R_{eff} we need to set the derivative of t'_{dFO4} with respect to R_{eff} equal to 0 and solve for R_{eff}. But first we must eliminate C_{G} from the expression for t'_{dFO4} since R_{eff} and C_{G} are connected through the relation $\tau' = R_{\text{eff}}C_{\text{G}}$, where $\tau'$is a constant. When we eliminate C_{G} we get:

$$t'_{\text{dFO4}} = 5\tau' + 24\, R_{\text{eff}}C + 36RC + 12\frac{R\tau'}{R_{\text{eff}}} \tag{12.100}$$

It is now clear that there are two terms that depend on R_{eff}. The derivative is:

$$\frac{d\, t'_{\text{dFO4}}}{d\, R_{\text{eff}}} = 24C - \frac{12R\,\tau'}{R_{\text{eff}}^2} \tag{12.101}$$

When we set the derivative equal to 0 and solve for R_{eff} we find:

$$R_{\text{eff}} = \sqrt{\frac{R\,\tau'}{2C}} \tag{12.102}$$

12.8 Layout

Solution 8.1 *Problem is on page 27.*

a) In the p-net graph all vertices have the order 2; so zero nodes with odd order. In the n-net graph there are two nodes with odd order: the output node and GND. So Euler paths are OK but we also have to check the order of the nodes. All paths in the n-net graph has to start with the D vertice. There are two possible orders that also work with the p-net, which is a ring: (D,C,B,A) and (D,A,B,C).

b) In the p-net there are two nodes with order three: VDD and the output node. In the p-net there are also two nodes with order three: GND and the node above the A2 and B2 transistor that are in parallel connected to GND. So for each of the nets it is possible to find a Euler path, but there are constraints on the start and end nodes for both paths.

Now we must check if it also possible to use the same order in both graphs. The limiting cases are often the nodes where there are no options, that is the ones connected in series. There are more of those in the n-net where we find four in row: A1,B1,B2,A2. In the n-net we could start at each end of this row since it is connected to the odd-vertice nodes at each end. However, when we check the p-net we find that there is only one possibility there and that is to start at VDD with A1. So we then conclude that there are two orders that work for both paths and they are: (A1, B1, B2 A2, A2, B2) and (A1, B1, B2 A2, B2, A2).

Solution 8.2 *Problem is on page 27.*

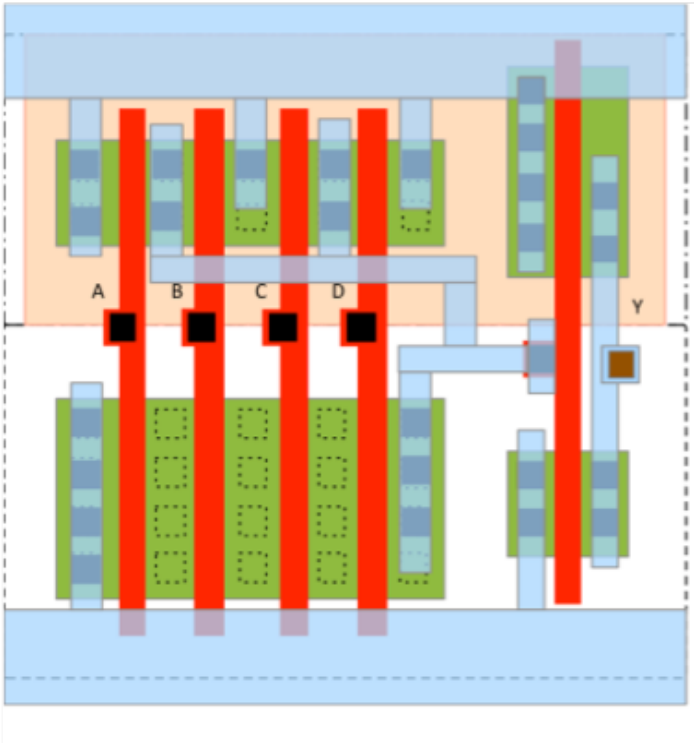


Figure 12.11: One possible layout of the AND4 gate.

1. Layout with minimal number of diffusion areas is shown in Figure 12.11.

2. For our layout of the 4-input NAND gate, with worst-case path resistance R in both n and p paths, there are two p-diffusion areas with width 2 connected to the output and 1 n-diffusion area with width 4. For a CMOS inverter with resistance R, there are diffusion areas of width 2+1 connected to the output (corresponding to $p_{inv} = 1$ since that was given in the problem). So $p_{NAND4} = (2+2+4)/3 = 8/3$ which is quite a bit less than 4 which is what we get from the schematic.

3. We need to find an expression for the normalized delay d of the AND4 gate as $d = p_{\text{AND4}} + g_{\text{AND4}} h_{\text{AND4}}$ where p_{AND4} is the part of the normalized delay that does no depend on the load capacitance and h_{AND4} is $\frac{C_L}{C_{\text{inAND4}}}$. For the entire gate the expression for the normalized delay is

$$d_{\text{AND4}} = g_{\text{NAND4}} h_{\text{NAND4}} + p_{\text{NAND4}} + h_{\text{inv}} + p_{\text{inv}}. \tag{12.103}$$

In this expression the only part that depends on external load capacitance is h_{inv}. The other three terms in the expression will form p_{AND4} so we have:

$$p_{\text{AND4}} = g_{\text{NAND4}} h_{\text{NAND4}} + p_{\text{NAND4}} + p_{\text{inv}}. \tag{12.104}$$

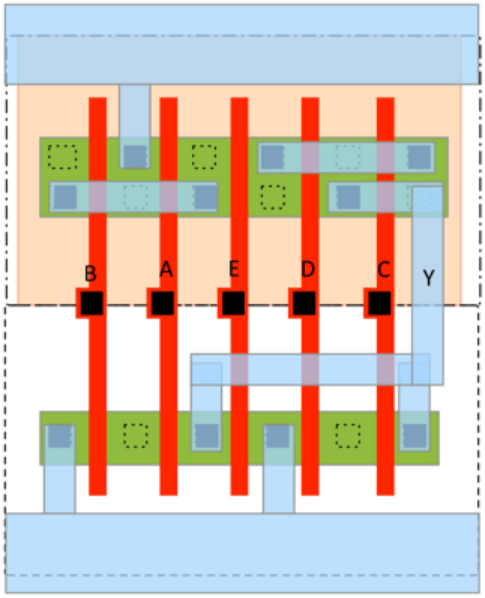
The electrical effort for the NAND4 gate, h_{NAND4} is 1, because its load capacitance is the same as its input capacitance. Thus, we find:

$$p_{\text{AND4}} = 2 \times 1 + \frac{8}{3} + 1 = 5\frac{2}{3} \tag{12.105}$$

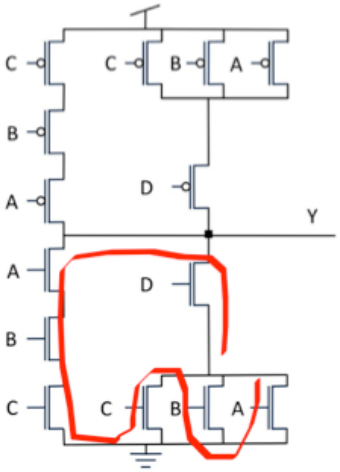
The input capacitance of the NAND4 gate is the same as for the inverter. Therefore, the electrical effort will be 1 also for the AND4 gate. In conclusion the solution is $p_{\text{AND4}} = 5.67$ and $g_{\text{AND4}} = 1$.

Solution 8.3 *Problem is on page 28.*

a) There are several solutions to this layout problem. In the figure below you see one possibility.

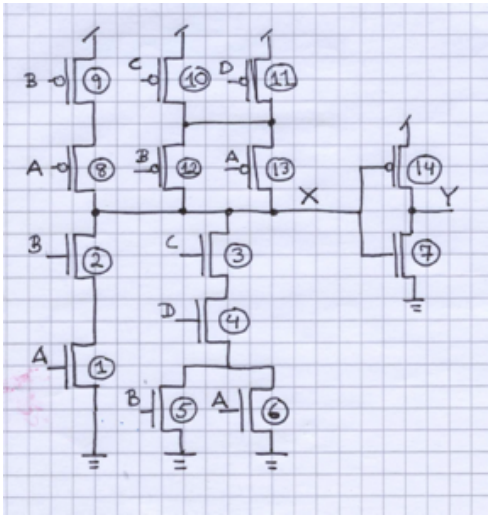


b) Below is the schematic with the order of the transistors shown. The logical function of the gate is $Y = \overline{ABC + D(A + B + C)}$. The reason the circuit is symmetrical is that the direct inverse of this function is $\overline{Y} = (A + B + C)(D + (ABC)) = D(A + B + C) + ABC$.



Solution 8.4 *Problem is on page 28.*

a) For reference we number the nMOS transistors in the layout 1-7 from left to right and the pMOS transistors 8-14 from left to right. We name the output of the compound gate X and the output of the inverter Y. The corresponding transistor schematics is shown below; the numbers of each transistor is to the right of that transistor:



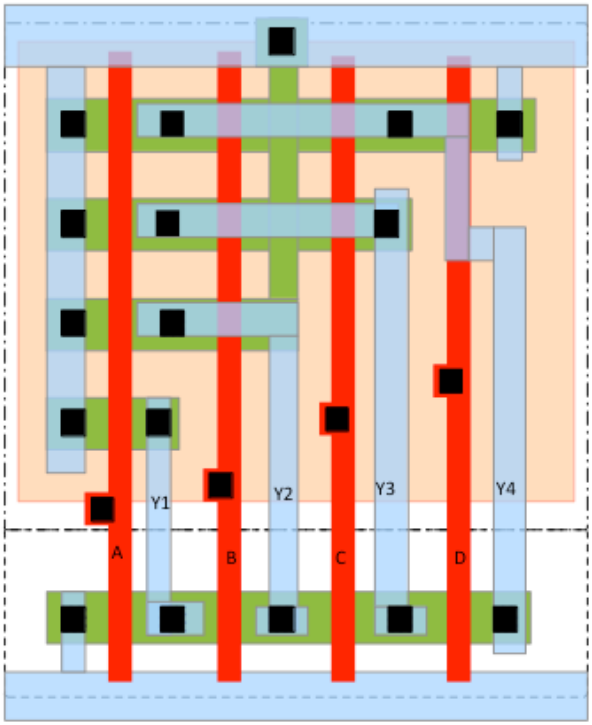
b) Solution:

$$Y = AB + (A + B)CD \tag{12.106}$$

It is easiest to find the function from the n-net of the compound gate. Since its output, X, is inverted to form Y, the n-net gives the function for Y directly.

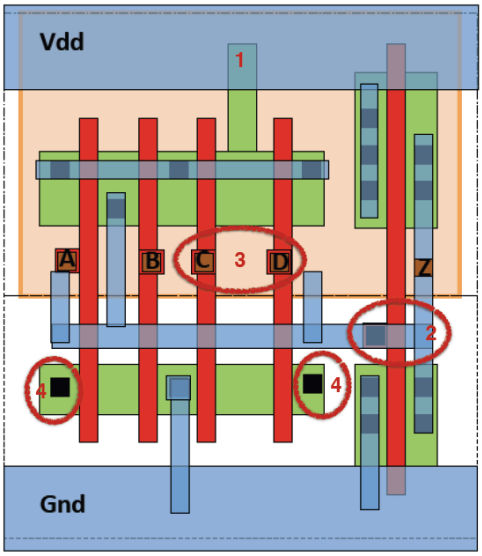
Solution 8.5 *Problem is on page 29.*

The layout is shown here:



Solution 8.6 *Problem is on page 29.*

The discrepancies are marked in the layout here:



The discrepancies are:

1. Missing contact from p-active to VDD metal-1.
2. Inverter input and output shorted.
3. Wrong order of inputs compared to schematics.
4. Accidentally misplaced metal-1 wires.

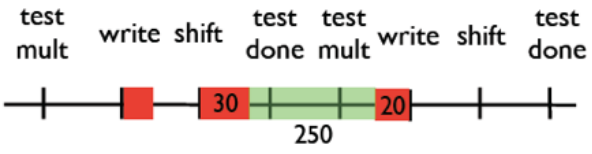
12.9 Sequential circuits

Solution 9.1 *Problem is on page 31.*

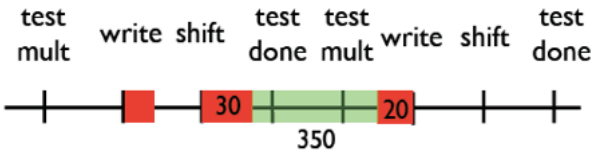
Solution 9.2 *Problem is on page 31.*

Solution 9.3 *Problem is on page 32.*

a) Both 16-bit adders have a worst-case delay of 250 ps. The figure below shows the unrolled steps in the control logic of the multiplier in the worst-case situation where every step results in a write to the product register. From the figure it is clear that three steps are completed in 300 ps. Thus, one step takes 100 ps and the resulting clock frequency is $1/(100\,\text{ps}) = 10\,\text{GHz}$. An 8-bit multiplication takes $1 + 8 \times 4 = 33$ steps which is then 3300 ps or 3.3 ns.



b) We will use the Sklansky adder because it will have a shorter worst-case delay. From the table, and our knowledge about prefix adders, we find that the expression for the worst-case delay is $50\,\text{ps} + \log 2(n) \times 50\,\text{ps}$. For n=64 we then have the worst-case delay as $50 + 6 \times 50 = 350$ ps. The ripple-carry adder has much longer delay since its delay grows linearly with n. The figure below show the unrolled steps in the control logic of the multiplier where it becomes clear that three steps are completed in 400 ps. Thus, one step takes 133 ps and the resulting clock frequency is $1/(133\,\text{ps}) = 7.5\,\text{GHz}$. A 32-bit multiplication with worst-case data lasts for $1 + 32 \times 4 = 129$ steps which each takes 133 ps. All in all then 17.2 ns.



c) See figure below from Hennessy and Patterson Computer Organization and Design. By shifting the product rather than the multiplicand we can use a 32-bit adder rather than a 64-bit adder. The 32-bit Sklansky adder has a worst-case delay of 300 ps. The cycle time is then 350/3 = 117 ps. The time for one multiplication to complete with worst-case data is then $129 \times 117\,\text{ps} = 15.1\,\text{ns}$. So the gain in calculation delay is not that great. One saves space and power too though.

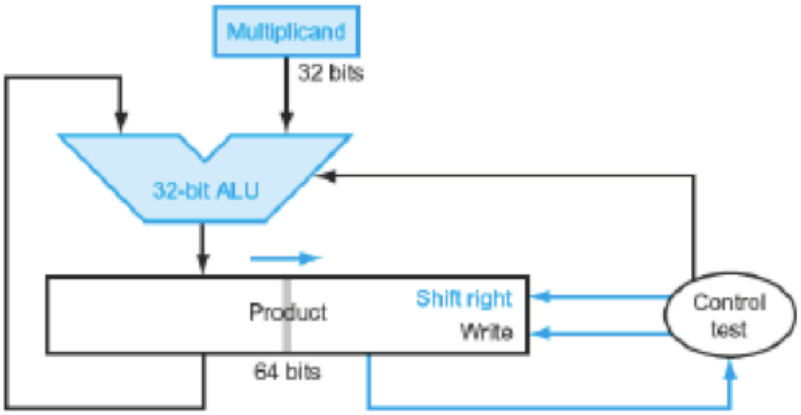


FIGURE 3.5 Refined version of the multiplication hardware. Compare with the first version in Figure 3.3. The Multiplicand register, ALU, and Multiplier register are all 32 bits wide, with only the Product register left at 64 bits. Now the product is shifted right. The separate Multiplier register also disappeared. The multiplier is placed instead in the right half of the Product register. These changes are highlighted in color. (The Product register should really be 65 bits to hold the carry out of the adder, but it's shown here as 64 bits to highlight the evolution from Figure 3.3.)

Solution 9.4 *Problem is on page 33.*

a) The propagation delay through the entire adder is:

   a. Full-adder 1: Max of propagation delays from A,B, and C_{in} inputs to C_{out} output

   b. Full-adder 2: Propagation delay from C_{in} to C_{out}

   c. Full adder 3: Max propagation delays from C_{in} to Sum and C_{out}

   With numbers we get: t_{pd} = max(25,20)+20+max(20,20)= 65 ps

   The scheduling overhead is $t_{\text{sched}} = t_{pcq} + t_{setup} = 35 + 30 = 65$ ps.

   All in all $T_c = t_{\text{pd}} + t_{\text{sched}} = 65 + 65 = 130\,\text{ps} \rightarrow f_{\text{clk}} = 7.7\,\text{GHz}$.

b) The minimum time until any output changes at the output of the adder is: t_{ccq} + minmimum of contamination delays from inputs A, B, C_{in} to Sum output for the full adder. With numbers we get: 21 [ps] + min(22, 15)[ps] = 36 ps. The change at the adder output is not allowed to happen within the hold time because then we have a hold violation. We have thold = 10 ps. So thus the maximum possible clock skew is:

$$T_{\text{skew}} \leq t_{ccq} + t_{cd,C_{\text{in}},C_{\text{out}}} - t_{\text{hold}} \tag{12.107}$$

$$T_{\text{skew}} \leq 21 + 15 - 10 = 26\,\text{ps} \tag{12.108}$$

c) Description: When we have the slow-slow and fast-fast corners the calculation for maximum clock frequency has to be repeated for the slow-slow corner only because all delays will be shorter for fast-fast corner. However, a hold violation can happen for any condition, so we have to check both corners when calculating the maximum allowed clock skew.

Calculation: For an update of the solution for task a) we arrive at these values from the slow-slow column in the table:

$$t_{\text{pd}} = \max(30, 25) + 25 + \max(25, 25) = 80\,\text{ps} \tag{12.109}$$

The scheduling overhead in the slow-slow corner is: tsched = pcq + tsetup = 40 + 35 = 75 ps All in all we find: $T_c = t_{\text{pd}} + t_{\text{sched}} = 80 + 75$ [ps] = 155 ps $\rightarrow$ $f_{\text{clk}} = 6.45$ GHz. For the solution in b) we have to check the requirement for both corners. In both cases we have $t_{cd,C_{\text{in}},C_{\text{out}}} < t_{cd,AB,C_{\text{out}}}$ so the requirement can still be expressed as $T_{\text{skew}} \leq t_{ccq} + t_{cd,C_{\text{in}},C_{\text{out}}} - t_{\text{hold}}$ for both corners:

$$\text{Fast-fast}: T_{\text{skew}} \leq 16 + 12 - 5 = 23\,\text{ps} \tag{12.110}$$

$$\text{Slow-slow}: T_{\text{skew}} \leq 24 + 20 - 20 = 24\,\text{ps} \tag{12.111}$$

All in all, taking the additional corners into account the maximum clock frequency is 6.45 GHZ and the maximum allowed clock skew is 23 ps.

12.10 Power, energy and scaling

Solution 10.1 *Problem is on page 35.*

For dynamic power the expression is:

$$P_{\text{dyn}} = \alpha f C_{\text{L}} V_{\text{DD}}^2 \tag{12.112}$$

Thus, decreasing V_{DD} decreases power (with quadratic dependence). For the static power we assume that the subthreshold current dominates. The expression for the power is

$$P_{\text{sub}} = V_{\text{DD}} \times I_{\text{sub}} \tag{12.113}$$

Here, V_{DD} decreases too, but the subthreshold current increases more due to the decrease in threshold voltage. Remember that a 100 mV decrease in threshold voltage causes a tenfold increase in the current. So the static power increases.

Solution 10.2 *Problem is on page 35.*

For dynamic power the expression is:

$$P_{\text{dyn}} = \alpha f C_{\text{L}} V_{\text{DD}}^2 \tag{12.114}$$

Thus, for dynamic power we have V_{DD}: **D**, V_{T} :**N**, C_{L}: **D** and Width: **N.**

For the short-circuit power we have we assume that $P_{\text{SC}} = V_{\text{DD}} \times I_{\text{SC}}$ and it is the regular drain-current equation in the saturation region that gives the short-circuit current. This current flows when both transistors are on during the transition, so we can roughly assume that the input voltage is $V_{\text{DD}}/2$ if we want to have a detailed expression for it:

$$P_{\text{SC}} = V_{\text{DD}} I_{\text{SC}} \approx V_{\text{DD}}\, k' \frac{W}{L} (\frac{V_{\text{DD}}}{2} - V_{\text{T}})^2 \tag{12.115}$$

Thus, for short-circuit power we have V_{DD}: **D**, V_{T}: **I** ,C_{L}: **N** and Width: **D.**

Similarly, for the static current the power is $P_{\text{sub}} = V_{\text{DD}} \times I_{\text{sub}}$. The subthreshold (leakage) current depends exponentially on the gate-to-source voltage; the further below the threshold voltage, V_{T}, V_{GS} is when the transistor is fully off, the lower the subthreshold current; so a higher threshold voltage gives a lower subtreshold current when $V_{\text{GS}} = 0$ V. Also again, the current is proportional to the transistor width.

Thus, for static power we also have V_{DD}: **D**, V_{T}: **I,** C_{L}: **N** and Width: **D.**

Solution 10.3 *Problem is on page 35.*

The reason for the difference is the stack effect. Both nMOS transistors are off when both A and B are ”0”. That is why the fist input entry in the table has the lowet current.

The best combinations for the carry cell is when all inputs are ”1” or all inputs are ”0”. In those cases all pMOS transistors or all nMOS transistors are OFF and we have the benefit of the stack effect.

Solution 10.4 *Problem is on page 36.*

c) The dynamic power consumption due to recharging of the capacitances is $P_{\text{dyn}} = \alpha f C_{\text{tot}} V^2$where α and f are given in the problem statement. V in the equation is V_{DD}. The remaining factor is C_{tot} , the total switched capacitance, which we have to determine. C_{in} for the inverter is not given in the problem but we can express the dynamic power in C_{in} . How many times Cin is the switched capacitance? It is

$$C_{\text{tot}} = C(1 + 4.5 + 4.5f + 4.5f_2 + 4.5f_3 + 4.5f_4) \tag{12.116}$$

if we count the also the input capacitance of the first inverter. We can also write the capacitance as

$$C_{\text{tot}} = C_{\text{parastictot}} + C_{\text{inputtot}} = 1/2 + 2 + 32 + 128 + (1 + 4 + 16 + 64 + 256 + 1024). \tag{12.117}$$

Either way the dynamic power is $0.25 \times 200\,\text{MHz} \times 1\,\text{V}^2 \times 1527.5 \times C_{\text{in}}$ where C_{in} should be a few fF. Even though we do not have the exact number for the input capacitance we can check that the numbers are reasonable using our previous knowledge of the 65 nm process. If we assume that C_{in} for the first inverter is a little less than 4 fF we can approximate $1527.5 \times C_{\text{in}}$ with 6000 fF = 6 pF. With this capacitance we arrive at

$$P_{\text{dyn}} = 50\,\text{MHzV}^2 \times 6\,\text{pF} = 300\,\mu\text{W} \tag{12.118}$$

which seems an entirely reasonable result.

d) It is incorrect because we neglect to take into account that if we have very large fanout for a gate the switching will be very slow at the gate output because its current is too small to charge the capacitance quickly. Then the n-net and p-net transistors in the gate will both be conducting at the same time and quite a large short-circuit current will flow during the switching.

Solution 10.5 *Problem is on page 36.*

a) (This task is also an example in Weste & Harris) The capacitance for the logic part is 50 million transistors is

$$50 \times 10^6 \times 0.3\,\mu\text{m} \times 1.8\,\text{fF}/\mu\text{m} = 27\,\text{nF} \tag{12.119}$$

and for the 950 million transistors in the memories:

$$950 \times 10^6 \times 0.1\,\mu\text{m} \times 1.8\,\text{fF}/\mu\text{m} = 171\,\text{nF}. \tag{12.120}$$

The power consumption for the logic part is then $0.1 \times 1\,\text{GHz} \times 1.44\,\text{V}^2 \times 27\,\text{nF} = 3.88\,\text{W}$. The dynamic power for the memories is: $0.02 \times 1\,\text{GHz} \times 1.44\,\text{V}^2 \times 171\,\text{nF} = 4.92\,\text{W}$. All in all the dynamic power for the chip is **8.8 W**.

b) (This task is also an example in Weste & Harris) The static power is due to the subthreshold leakage and gate leakage. We assume that half of all transistors as on and half are off. Only off transistors contribute subthreshold leakage and only on transistors contribute gate leakage current. For the memory part we have the total leakage current:

$$425 \times 10^6 \times 0.1\,\mu\text{m} \times (10\,\text{nA}/\mu\text{m} + 5\,\text{nA}/\mu\text{m}) = \mathbf{637\,mA} \tag{12.121}$$

For the logic part we have 25 million transistors that are on, and 25 that are off. Of the off ones 5 % have the high VT. So the leakage current is:

$$25 \times 10^6 \times 0.3\,\mu\text{m} \times (0.95 \times 10\,\text{nA}/\mu\text{m} + 0.05 \times 100\,\text{nA}/\mu\text{m} + 5\,\text{nA}/\mu\text{m}) = \mathbf{146.25\,mA}. \tag{12.122}$$

All in all the leakage current is 783.75 mA and the power (since $P = U \times I$) is **940 mW**. (Note that this result is not the same as in the book because there is a calculation error in book solution).

c) The dynamic power for the logic part would increase by 20 % since the capacitance increases by 20 % and all the other factors stay the same. The leakage current would be

$$0.3\,\mu\text{m} \times 10^6 \times ((0.99 \times 25 + 5) \times 10\,\text{nA}/\mu\text{m} + 0.01 \times 25 \times 100\,\text{nA}/\mu\text{m} + 30 \times 5\,\text{nA}/\mu\text{m}) = \mathbf{270\,mA}. \tag{12.123}$$

So it would not save any dynamic power because the added 5 million transistors have more leakage than what we save by having fewer low VT transistors.

d) The dynamic power of 3.88 W from a) corresponds to 3.23 A of current. When this current is drawn through the power-gate switch there should be no more than 60 mV of voltage drop across it. Using Ohm's law, we find the maximum resistance of $R = 0.06\,\text{V}/3.23\,\text{A} = 0.0186\,\Omega$. So the transistor has to be very wide! $W = 2000\,\Omega\,\mu\text{m}/0.0186\,\Omega = 107\,526\,\mu\text{m} =$ **108 mm**. So the transistor is around 11 cm wide! (In practice it can be a bit less wide since the transistor resistance at low V_{DS} is smaller than R).

e) The capacitance of the switch is $W \times 1\,\text{fF}/\mu\text{m}$; that results in 107 526 fF or 107.5 nF. The energy is then 155 nJ (since $E = CV_{DD}^2$ and V_{DD} is 1.2 V). The static power for the logic part from b) is 146 mA × 1.2 V = 175.2 mW. Power is energy per time. So how long time for the total energy due to leakage to be equal to E_{sw}? We get $E_{sw} = E_{leak} = P_{leak} \times t$ so $t = E_{sw}/P_{leak} = 186\,\text{nJ}/175.2\,\text{mJ s}^{-1}$ (since Watts are Joules/second). We finally arrive at $t =$ **0.89 µs**.

f) BONUS QUESTION Obviously the case in c) is not good since both the dynamic power and the leakage is higher than in b), but in the general case it is a question about determining when to spend all the energy to turn the power off. One has to be rather good at predicting the down-time to spend the energy required. So it may be better to spend more on the dynamic power if the static power can be reduced without having to turn the power off, since it is very costly to do so.

Solution 10.6 *Problem is on page 37.*

We have to redo the calculations for tasks a)-b) and d)-e) in solution 12.51 but with $V_{DD} = 1$ V.

a) Capacitance remains the same, of course. So the power consumption for the logic part is then $0.1 \times 1\,\text{GHz} \times 1\,\text{V}^2 \times 27\,\text{nF} = 2.7\,\text{W}$. The dynamic power for the memories is: $0.02 \times 1\,\text{GHz} \times 1\,\text{V}^2 \times 171\,\text{nF} = 3.42\,\text{W}$. All in all the dynamic power for the chip is **6.12 W**.

b) We have the same leakage currect as before, if we assume that the numbers in the problem statement holds also at the lower V_{DD}. All in all the leakage current is still 783.75 mA and the power is (since $P = U \times I$) **784 mW**.

d) The dynamic power of 2.7 W from a) corresponds to 2.7 A of current. When this current is drawn through the power-gate switch there should be no more than 50 mV of voltage drop across it. Using Ohm's law, we find the maximum resistance of $R = 0.05\,\text{V}/2.7\,\text{A} = 0.0186\,\Omega$. So the transistor has to be very wide! $W = 2000\,\Omega\,\mu\text{m}/0.0186\,\Omega = 107\,526\,\mu\text{m} =$ **108 mm**. So the transistor is around 11 cm wide! (In practice it can be a bit less wide since the transistor resistance at low V_{DS} is smaller than R).

e) The capacitance of the switch is $W \times 1\,\text{fF}/\mu\text{m}$; that still results in 107 526 fF or 107.5 nF. The energy is however lower: 107.5 nJ (since $E = CV_{DD}^2$ and V_{DD} is 1.0 V). The static power for the logic part from b) is now 146 mA × 1.0 V = 146 mW. Power is energy per time. So how long time for the total energy due to leakage to be equal to E_{sw}? We get $E_{sw} = E_{leak} = P_{leak} \times t$ so $t = E_{sw}/P_{leak} = 107.5\,\text{nJ}/146\,\text{mJ s}^{-1}$ (since Watts are Joules/second). We finally arrive at $t =$ **0.74 µs**. So it is a bit shorter than at 1.2 V, even though the allowed voltage drop is smaller here.

Solution 10.7 *Problem is on page 37.*

Preliminaries

Application: The digital video in the video-rendering application has 25 frames per second. Thus the maximum time for the calculations for one frame is 40 ms. One frame is 640 × 480 pixels, that is 307 200 pixels.

PP processor: The PP processor requires 3 072 000 clock cycles for the calculations for one frame.

a) The first task is to fill in the empty cells in Table 10.2. To find the maximum clock frequency for lower V_{DD} we need to know how the delay (that is τ) in the process scales with V_{DD}. We know that $\tau = 0.7RC$. In this expression only R changes with V_{DD} . We know that $R = \frac{V_{DD}}{I_{DSAT}}$ and that $I_{DSAT} \sim (V_{DD} - V_T)^2$ if we assume that the quadratic current equations hold as stated in the problem. Thus, the ratio of the max clock frequencies at two supply voltages can be written as:

$$\frac{f_{clk2}}{f_{clk1}} \sim \frac{\tau_1}{\tau_2} \sim \frac{\frac{V_{DD1}}{I_{DSAT1}}}{\frac{V_{DD2}}{I_{DSAT2}}} \sim \frac{\frac{V_{DD1}}{(V_{DD1}-V_T)^2}}{\frac{V_{DD2}}{(V_{DD2}-V_T)^2}} \tag{12.124}$$

For our three supply voltages $\frac{V_{DD}}{(V_{DD}-V_T)^2}$ evaluates to: for 1.2 V: 1.48, for 1.0 V: 2.04, and for 0.8 V: 3.2.

Thus, we have $\frac{f_{clk1.0}}{f_{clk1.2}} = \frac{1.48}{2.04} = 0.73$ and $\frac{f_{clk0.8}}{f_{clk1.2}} = \frac{1.48}{3.2} = 0.46$.

The current corresponding to the dynamic power consumption at the maximum clock frequency is

$$f_{clkmax}\alpha C V_{DD}. \tag{12.125}$$

(It has to be multiplied by the voltage to get the dynamic power.) So it scales with the maximum clock frequency (calculated above) and the supply voltage (the activity factor and capacitance does not change) as

$$\frac{I_{dyn2}}{I_{dyn1}} \sim \frac{f_{clk2}V_{DD2}}{f_{clk1}V_{DD1}} \tag{12.126}$$

Thus we have $\frac{I_{dyn1.0}}{I_{dyn1.2}} = 0.73\frac{1.0}{1.2} = 0.608$ and $\frac{I_{dyn0.8}}{I_{dyn1.2}} = 0.46\frac{0.8}{1.2} = 0.306$. The resulting current values are shown in Table 12.1. (4 p)

Table 12.1: Data for the PP processor

Supply voltage V_{DD} [V]	Maximum clock frequency [GHz]	Current due to dynamic power consumption @ max clock frequency and a realistic activity factor [mA]	Idle current @ room temperature @ max clock frequency and a low activity factor [mA]	Static current in sleep mode @ room temperature (clock signal turned off for logic, but clock generation maintained) [mA]	Static current in hibernation mode (clock generation stopped and internal supply voltages turned off) [µA]
1.2	1.0	600	100	60	60
1.0	**0.73**	**438**	80	37.5	60
0.8	**0.46**	**184**	64	28	60

Sleep mode: The time it takes to enter sleep mode is 10 µs and it takes 20 µs for the processor to wake up from sleep mode. The energy required to switch the clocks off is 10 µJ. **Hibernation mode**: The time it takes to enter hibernation mode is 1 ms and it takes 19 ms to wake up from hibernation mode. The energy required to turn off V_{DD} is 500 µJ.

b) It takes the same number of cycles to do the calculations at both voltages, but the cycle time differs. At 1 GHz the cycle time is 1 ns. At 460 MHz the cycle time is 2.17 ns. The energy is power times time. Thus, we arrive at: $E_{dyn} = I_{dyn} \times V_{DD} \times t_{cycle} \times N_{cycles}$. And we have for 1.2 V and 0.8 V:

$$E_{dyn1.2} = 600\,\text{mA} \times 1.2\,\text{V} \times 1\,\text{ns} \times 3.07 \times 10^6 = 2210\,\mu\text{J}$$
$$E_{dyn0.8} = 184\,\text{mA} \times 0.8\,\text{V} \times 2.17\,\text{ns} \times 3.07 \times 10^6 = 918\,\mu\text{J}$$

so a large reduction of the dynamic energy at the price of a more than doubled computations time.

c) The time in sleep mode is the time left after the computations for a frame are done. At 1.2 V the necessary computations take 3 ms and at 0.8 V they take 8 ms (we neglect the time it takes to switch to sleep mode since it is so short in comparison). The remaining times for a frame are 40 – 3 = 37 ms and 40-8 = 32 ms respectively. Thus, we have:

$$E_{sleep1.2} = 60\,\text{mA} \times 1.2\,\text{V} \times 37\,\text{ms} = 2664\,\mu\text{J}$$
$$E_{sleep0.8} = 28\,\text{mA} \times 0.8\,\text{V} \times 32\,\text{ms} = 717\,\mu\text{J}$$

d) Since we have so much time available for the computations for one frame we should run at the lowest possible supply voltage, in this case 0.8 V, to save energy. At that supply voltage the energy required to switch off the power supply is larger than what we could save during the 12 ms when the power supply would be off. So the answer is NO. We should not use hibernation mode for this application. (2 p)

Solution 10.8 *Problem is on page 38.*

Solution 10.9 *Problem is on page 39.*

See solution in Table 12.2.

Solution 10.10 *Problem is on page 40.*

See solution in Table 12.2.

Table 12.2: Table for Dennard scaling and constant voltage scaling.

Parameter	Sensitivity expression	Dennard , scaling factor S	Constant voltage, scaling factor S
Scaling parameters			
L: length		$1/S$	$1/S$
W: width		$1/S$	$1/S$
t_{ox}: gate oxide thickness		$1/S$	$1/S$
V_{DD}: power supply voltage		$1/S$	1
V_T : threshold voltage(s)		$1/S$	1
NA : substrate doping		S	S
Device characteristics			
β: current factor	$\frac{W}{L}\frac{1}{t_{ox}}$	S	S
I_{DS}: transistor current	$\beta(V_{DD}-V_T)^2$	$1/S$	S
R_{eff} resistance	$\frac{V_{DD}}{I_{DS}}$	1	$1/S$
C: gate capacitance	$\frac{WL}{t_{ox}}$	$1/S$	$1/S$
τ: gate delay	$R_{eff}C$	$1/S$	$1/S^2$
f: clock frequency	$\frac{1}{\tau}$	S	S^2
E: switching energy (per gate)	CV_{DD}^2	$1/S^3$	$1/S$
P: switching power (per gate)	Ef	$1/S^2$	S
A: area (per gate)	WL	$1/S^2$	$1/S^2$
Switching power density	$\frac{P}{A}$	1	S^3
Switching current density	$\frac{I_{DS}}{A}$	S	S^3

Solution 10.11 *Problem is on page 40.*

The FO4 delay is $(4 + p_{inv})\tau$. With no major changes in the transistor parasitics it should scale just as τ. If nothing else is said we assume Dennard scaling where "all" parameters are scaled the same. See table 12.2. In that table we also already derived the scaling of τ. We can redo it though: We have

$$\tau = RC = \frac{V_{DD}C}{I_{DSAT}} = \frac{V_{DD}C}{\frac{W}{L}\mu C_{\mathrm{OX}}(V_{\mathrm{DD}} - V_{\mathrm{T}})^2} \tag{12.127}$$

The only tricky part is μC_{OX} where we have to remember that it scales as $\frac{1}{t_{\mathrm{OX}}}$, the inverse of the gate-oxide thickness. Now we can derive the scaling:

$$\tau \sim \frac{\frac{1}{S} \times \frac{S}{S^2}}{\frac{S}{S} \times S \times \frac{1}{S^2}} = \frac{1}{S}. \tag{12.128}$$

The scaling of the transistor lengths is $S = \frac{0.35}{0.13} = 2.7$ so the new FO4 delay should be $\frac{125}{2.7} = 46\,\mathrm{ps}$. Here, however the V_{DD} is not scaled down with the full scaling, which would have resulted in a V_{DD} of 1.22 V. The ratio for V_{DD} is only: $K = \frac{3.3}{1.8} = 1.83$. We can derive the scaling for τ when the voltages are not scaled the same as as the transistors as:

$$\tau \sim \frac{\frac{1}{K} \times \frac{S}{S^2}}{\frac{S}{S} \times S \times \frac{1}{K^2}} = \frac{K}{S^2}. \tag{12.129}$$

As a result we expect the FO4 delay in the 0.13 μm process to be

$$\mathrm{FO4}_{0.13} = 125\,\mathrm{ps} \times \frac{1.83}{2.7^2} = 32\,\mathrm{ps}. \tag{12.130}$$

Solution 10.12 *Problem is on page 40.*

a) The specification $f = \alpha V_{\mathrm{DD}}$ means we have $P_{\mathrm{dyn}} = C \times V_{DD}^3$. We also know that the dual-core will have twice the capacitance of the single core. So then $P_{DUAL-CORE} = P_{SINGLE-CORE}$ means

$$2C(xV_{\mathrm{DD}})^3 = CV_{\mathrm{DD}}^3, \tag{12.131}$$

where x is the scale factor. Hence we get $x = \frac{1}{\sqrt[3]{2}} \approx \frac{1}{1.26}$. Thus, the new supply voltage has to be $V_{\mathrm{DD}} = \frac{1.2}{1.26} \approx 0.95\,\mathrm{V}$.

b) The scaling is from 0.90 to 0.65 µm, that is approximately $S = \frac{1}{\sqrt{2}}$. Thus the area will be half of the old one, i.e. $A_{\mathrm{new}} = \frac{A_{\mathrm{old}}}{2} = 100\,\mathrm{mm}^2$.

If frequency scales as V_{DD} it becomes $3.8/1.2 = 3.2\,\mathrm{GHz}$. Power: the capacitance is scaled to half, so the new power is $P_{\mathrm{dyn}_{\mathrm{new}}} == \frac{P_{\mathrm{dyn}_{\mathrm{old}}}}{2}(\frac{1}{1.2})^3 = \frac{100}{2}(\frac{1}{1.2})^3 \approx 30\,\mathrm{W}$.

c) We now assume 10 W of static power and 90 W of dynamic power to begin with. 10 W of static power at $V_{\mathrm{DD}} = 1.2\,\mathrm{V}$ corresponds to a leakage current of 8.3 A. A four-fold increase of the static leakage current would mean a 33 A leakage current, and hence, 33 W of static power at $V_{\mathrm{DD}} = 1\,\mathrm{V}$. The dynamic power was originally 90 W, and becomes $0.30 \times 90 = 27\,\mathrm{W}$ in the new technology, when we use the reduction factor we already calculated in task b). 55-45! The new total power dissipation is 60 W. And discouragingly enough, more static power than dynamic power is dissipated in this what-if scenario. Not good. Something has to be done to reduce leakage!

Solution 10.13 *Problem is on page 40.*

The propagation delay in a processor is given by

$$t_d = 0.7RC = 0.7\frac{V_{\mathrm{DD}}C}{I_{\mathrm{DSAT}}} \tag{12.132}$$

where the saturation current is given by the square-law model:

$$I_{\mathrm{DSAT}} = \frac{k}{2}(V_{\mathrm{DD}} - V_{\mathrm{T}})^2 \tag{12.133}$$

With four cores we can allow an increase in delay by a factor four. This increase in turn allows us to decrease the supply voltage, V_{DD}. To find the maximum allowed scale factor x we equate four times the expression for the original t_d with the expression for the new t_d. Remembering that V_{T} does not scale, we get:

$$4 \times \frac{1}{(1-0.25)^2} = \frac{x}{(x-0.25)^2} \tag{12.134}$$

which yields

$$x = 0.3125 + \sqrt{0.3125^2 - 0.06252} = 0.48 \approx \frac{1}{2} \tag{12.135}$$

So the new V_{DD} is half of the original one. Dynamic power dissipation is given by:

$$P_{dyn} = \alpha f C V_{\mathrm{DD}}^2 \tag{12.136}$$

If we assume that the activity factor α is the same in both cases, the power dissipation for the multicore solution is:

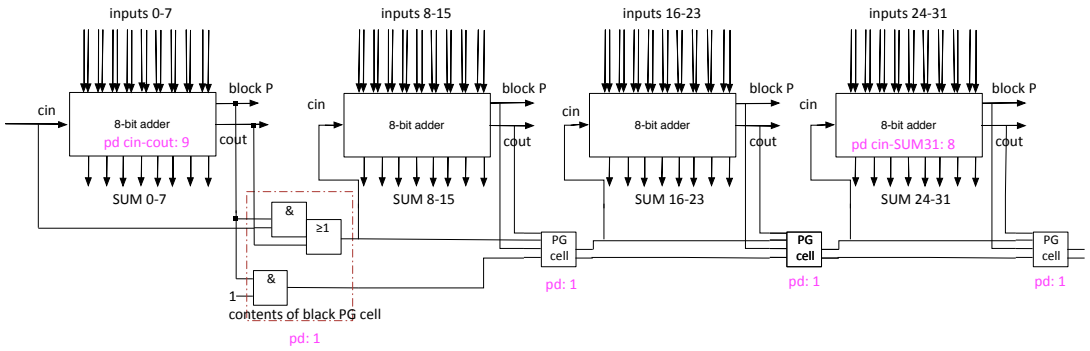
$$P_{dynmc} = \alpha \left(\frac{f}{4}\right) 4.2C \left(\frac{}{V}_{\mathrm{DD}} 2\right)^2 = \frac{4.2}{16} P_{dyn} \approx 0.26 P_{dyn}, \tag{12.137}$$

where we have increased the capacitance another $0.2C$ to account for the necessary wiring. So in conclusion the quad-core power dissipation is only a quarter of the single-core power dissipation. At least in the ideal world!

12.11 Adders

Solution 11.1 *Problem is on page 41.*

a) Here is the solution:



b) As hinted in the solution under a) the critical path is from c_{in} to c_{out} of the first 8-bit adder, through the 3 PG cells and from c_{in} to SUM31 of the last 8-bit adder. The delay to SUM31 is then: $9 + 3 + 8 = 20$ unit delays. We should also check that the delay to to the P and G outputs are shorter. They are both $9 + 3 + 1 = 14$ unit delays.

c) A similar solution with 4-bit adders would have shorter delays to generate the first carry out and the last SUM output, but twice the number of PG cells. The delay to the last SUM bit would then be: $5 + 7 + 4 = 16$ unit delays. So in this case it would pay off to use a shorter adder.

Solution 11.2 *Problem is on page 42.*

Solution shown in Figure 12.12

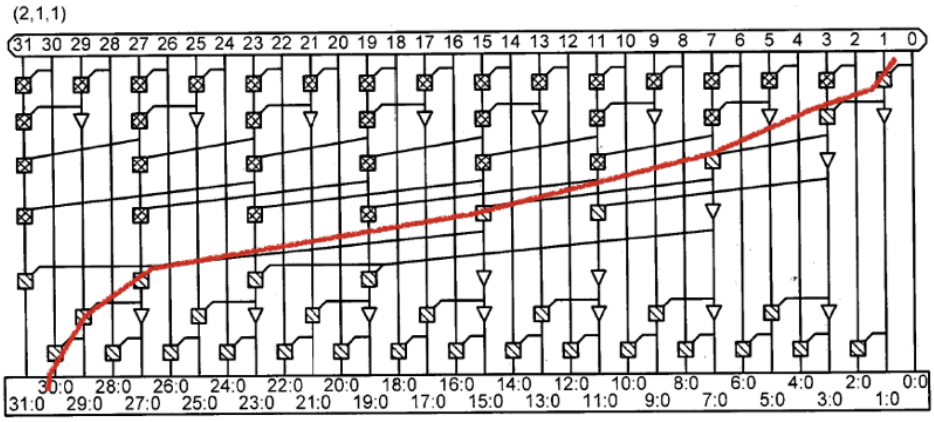


Figure 12.12: The unknown adder completed and with the critical path drawn.

Appendices

Appendix A

Templates and graphs to draw on

We placed all the large templates and graphs that you can draw on yourself in this appendix, so that if you want to print only these pages, that can easily be achieved without printing everything else. And conversely, if you want to print the exercises you do not get many pages just with the templates and graphs.

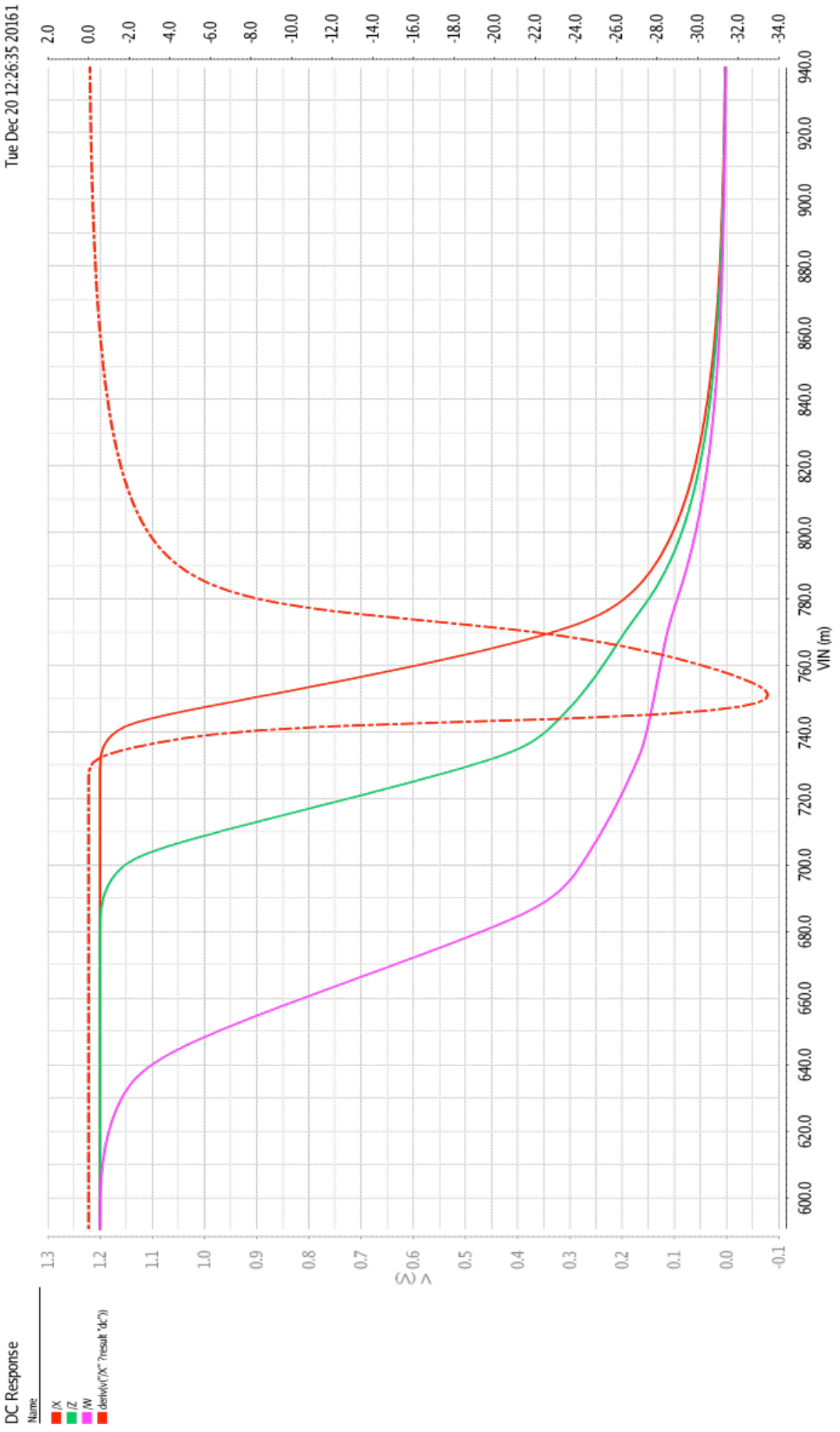


Figure A.1: Voltage transfer curves (VTC) for the three outputs X, Y and W and the derivative of the VTC for output X in larger scale.

Bibliography